

AWARD NUMBER: W81XWH-13-1-0284

TITLE: Biological and Clinical Characterization of Novel lncRNAs Associated with Metastatic Prostate Cancer

PRINCIPAL INVESTIGATOR: Dr. Rohit Malik

CONTRACTING ORGANIZATION: Regents of the University of Michigan
Ann Arbor, Michigan 48109-1274

REPORT DATE: October 2014

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE October 2014		2. REPORT TYPE Annual		3. DATES COVERED 15 Aug 2013 - 14 Aug 2014	
4. TITLE AND SUBTITLE Biological and Clinical Characterization of Novel lncRNAs Associated with Metastatic Prostate Cancer				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER W81XWH-13-1-0284	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Rohit Malik, Ph.D. E-Mail: romalik@med.umich.edu				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Regents of the University of Michigan 3003 S. State Street Ann Arbor, MI 48109-1274				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT lncRNA are polyadenylated RNA species that are more than 200bp in length and are implicated in development of variety of cancers including prostate cancer. In this study we have identified <i>PCAT29</i> as the first androgen receptor-repressed lncRNA that functions as a tumor suppressor. We show that loss of PCAT29 may identify a subset of patients at higher risk for disease recurrence. Collectively, PCAT29 is a tumor-suppressive lncRNA of prognostic and therapeutic relevance in prostate cancer. In addition, we have performed transcriptome analysis to identify a set of AR regulated lncRNAs. AR plays a critical role in the development and progression of prostate cancer. AR regulates a large repertoire of genes; however, its role in Long non-coding RNAs (lncRNA) regulation remains unclear. With this compendium of AR regulated lncRNA, we will be able to identify lncRNAs that play important role in prostate cancer development and progression.					
15. SUBJECT TERMS Prostate Cancer, Androgen receptor, long non-coding RNA					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	49	19b. TELEPHONE NUMBER (include area code)

Table of Contents

	<u>Page</u>
1. Introduction.....	1
2. Keywords.....	1
3. Accomplishments.....	2-11
4. Impact.....	11-12
5. Changes/Problems.....	12-13
6. Products.....	13-14
7. Participants & Other Collaborating Organizations.....	14
8. Special Reporting Requirements.....	15
9. Appendices.....	

INTRODUCTION

Despite improvements in medical treatments over the past three decades, prostate cancer remains the second most common cause of cancer death among U.S. men. According to National Cancer Institute, in 2012 ~241,000 prostate cancer diagnosis and ~28,000 related deaths of American men are estimated. Androgen deprivation, surgery, and/or radiotherapy in combination with chemotherapy has proven to be effective in treating patients that display localized disease; however, progression to hormone refractory aggressive disease in a subset of prostate cancer patients remains the primary cause of mortality. The molecular mechanisms that contribute to the progression of localized disease into an aggressive disease remain largely unknown. Long non-coding RNAs (lncRNAs), have recently emerged as key players in tumor biology and can potentially serve as promising biomarkers. However, the vast majority of lncRNAs remain undiscovered or uncharacterized entities and their roles in prostate cancer are unclear. Recent advances by our lab using high-throughput sequencing of prostate cancer tissues have led to the systematic identification of lncRNAs aberrantly expressed in prostate cancer that may play a role in prostate cancer progression. The two main objectives of this proposal were:

- 1) To evaluate the role of prostate lncRNAs in the progression of prostate cancer and,*
- 2) To explore the potential of prostate-specific lncRNAs to reliably predict disease progression.*

1. KEYWORDS

lncRNA, AR, PCAT, SCHLAP1, DHT, RNA

2. ACCOMPLISHMENTS

What were the major goals of the project?

Following three aims were proposed.

Specific Aim-1: To characterize nominated lncRNAs associated with metastatic castration-resistant prostate cancer. (Month 1-18)

This specific aim is 100% completed

Specific Aim 2: To elucidate the function of lncRNA in metastatic castrate resistant prostate cancer: (Month 6-24)

This aim is 90% completed

Specific Aim 3: To identify lncRNAs that serve as potential biomarkers of disease progression. (Month 12-24)

This aim is 50% completed

What was accomplished under these goals?

Various tasks proposed in individual specific aims are written below. Results obtained under each task are shown in form of figures.

Specific Aim-1: To characterize nominated lncRNAs associated with metastatic castration-resistant prostate cancer. (Month 1-18):

Task 1: To validate predicted transcripts in prostate cancer cell lines by qPCR (Month 1-3)

Progress (100% completed)

We validated several lncRNAs including: PCAT29, SCHLAP, G5303 and PCAT51. The expression of each lncRNA in prostate cell lines was determined by qPCR. Expression of PCAT29 and SCHLAP-1 has been published (see appendix)

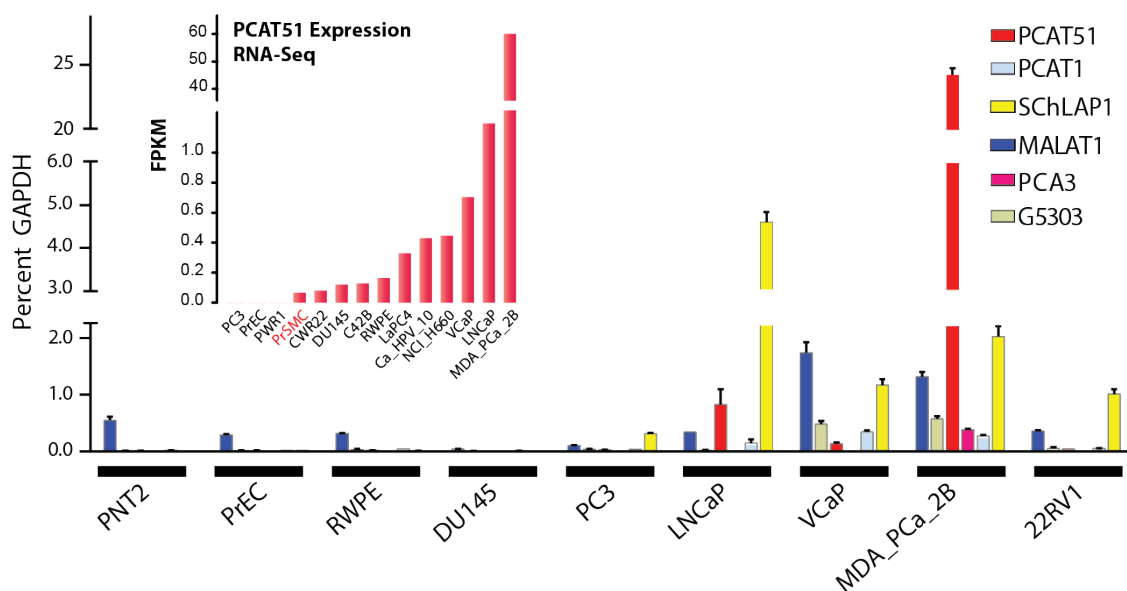


Figure1: Expression of each lncRNA in prostate cancer cell lines. Inset shows the expression of PCAT51 determined by RNA-seq in prostate cancer cell lines.

Task 2: To determine exon structure by rapid amplification of cDNA ends (RACE) (Month 1-6):

Progress (100% completed)

We choose PCAT29, SCHLAP and PCAT51 for further study. We performed RACE on various cell lines to determine the exon structure of each transcript. Show below is exon structure of PCAT51 as determined by RACE. Further, Northern blotting was performed to confirm the size of PCAT51. Exon structure of PCAT29 and SCHLAP are published (see appendix)

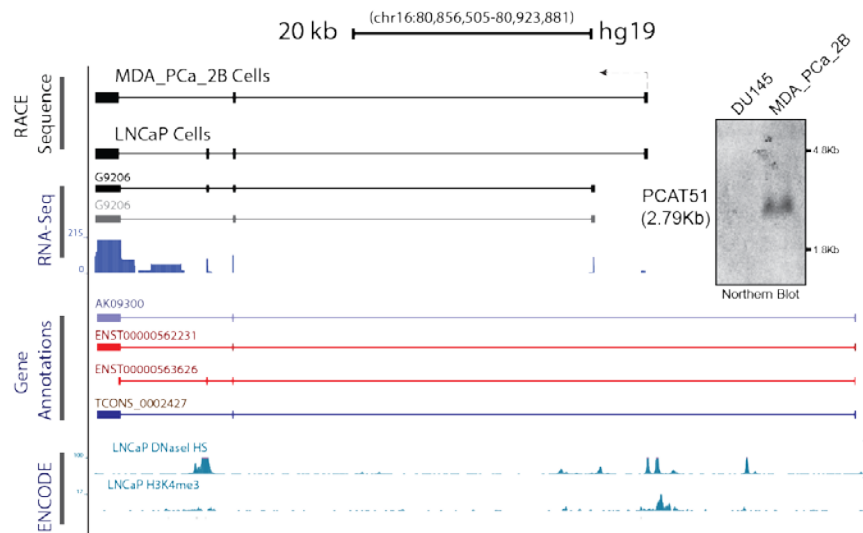


Figure 2: Genome browser view of PCAT51. Exon structure determined by RACE as well as other gene annotation is shown. Inset shows PCAT51 northern blot.

Task 3: To construct siRNA for the most conserved isoform and screen for lncRNA knockdown efficiency of by qPCR (Month 6-9)

Progress (100% completed)

We made 4 siRNA targeting PCAT51. These siRNA were tested in LNCaP and MDA_PCa_2b cells. In both the cell lines we got 80% knockdown of the gene. Gene expression was quantified by qRT-PCR. siRNA for PCAT29 and SCHLAP are now published (see appendix)

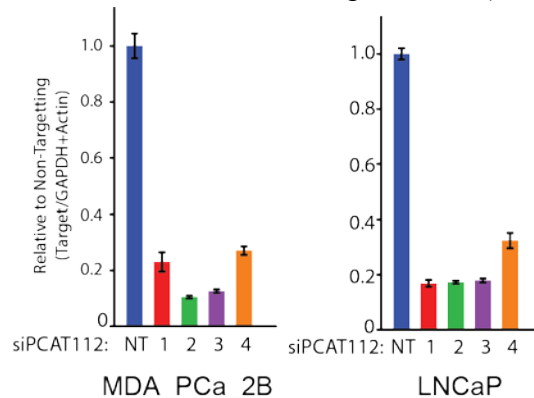


Figure 3: MDA_PCa_2B and LNCaP cells were transfected with 4 independent siRNAs targeting PCAT51. Expression of PCAT51 was determined by q-RT-PCR

Task 4: To generate stable knockdown cell lines using shRNA constructs (Month 6-9):

Progress (100% completed)

We made shRNA constructs targeting PCAT51, PCAT29 and SCHLAP1. Two best siRNA sequences targeting the genes were used. shRNA knockdown data for PCAT29 and SCHLAP1 is published (See appendix).

Task 5: To clone two most abundant lncRNA isoforms into pcDNA vector (month 6-9)

Progress (100% completed)

Two most abundant isoforms of PCAT29, SCHLAP1 and PCAT51 were amplified by PCR. These isoforms were then PCR amplified and cloned into lentiviral vector (pcDH).

Task 6: To generate stable cell lines (RWEP) over-expressing lncRNAs by lentiviral transduction (Month 6-12)

Progress (100% completed)

Viral particles were made following generation of stable cell lines expression individual lncRNAs. These cells lines were tested for gene overexpression.

Task 7: To perform *in-vitro* proliferation, migration and invasion assays using stable cell lines (Month 9-12)

Progress (100% completed)

We performed *in-vitro* proliferation, migration and invasion assays on PCAT29 KD and overexpression cells (Figure 4). As shown below, KD of PCAT29 increased cell proliferation and migration and overexpression of PCAT29 decreased cell proliferation and migration suggesting a tumor suppressor role for PCAT29.

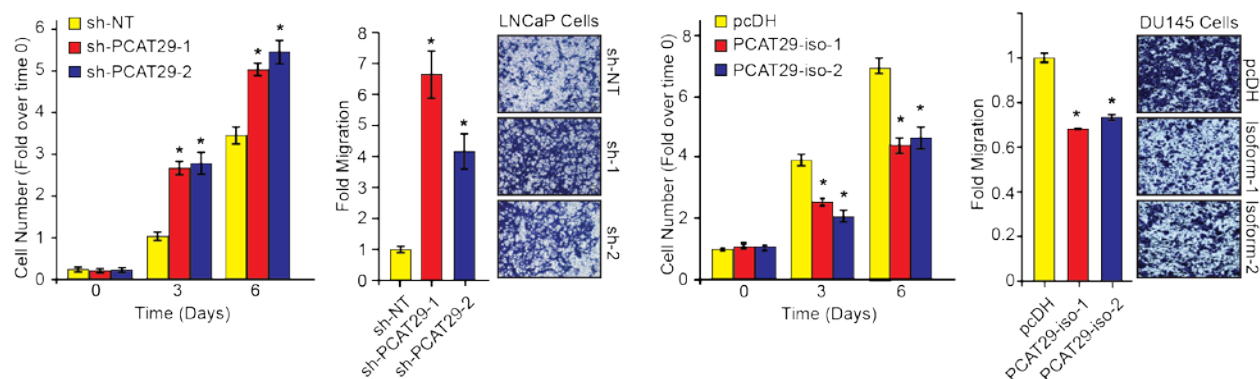


Figure 4: Proliferation and migration of LNCaP cells stably expressing PCAT29 shRNA and DU145 cells expressing PCAT29 expression constructs. Representative micrographs of crystal violet–stained migrated cells are shown as insets. Data, mean \pm SEM. *, $P < 0.05$ by the Student t test.

Similar to PCAT29, we performed proliferation and apoptosis assay on PCAT51 KD cells. As shown below PCAT51 KD using siRNA reduced cell proliferation and induced apoptosis suggesting an oncogenic role for PCAT51. Results for SCHLAP are now published (see appendix)

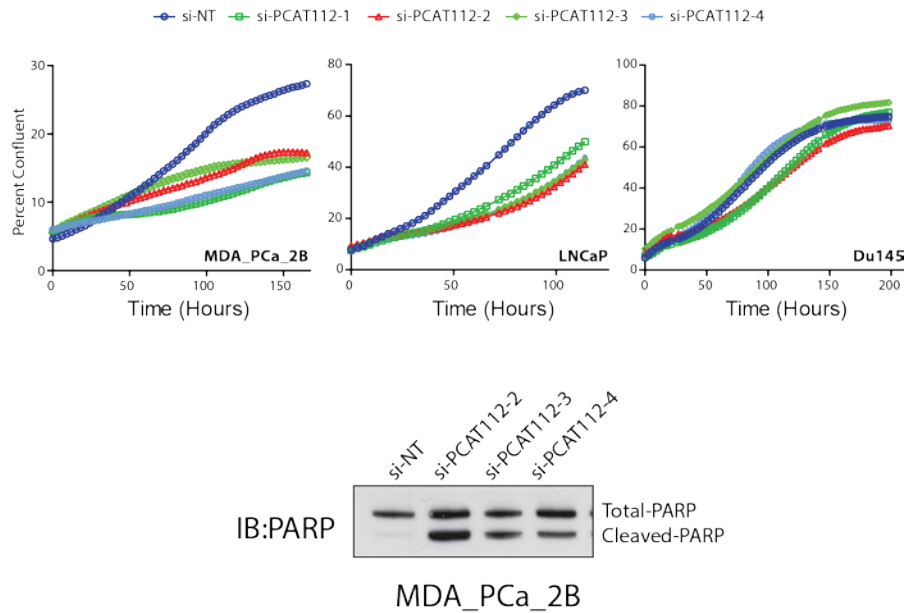


Figure 5: Proliferation of MDA_PCa_2B, LNCaP and DU145 cells transfected with siRNA targeting PCAT51. Effect on apoptosis was determined by immunoblotting for cleaved PARP

Task 8: To perform *in-vivo* tumor formation and metastasis assay using stable cell lines in collaboration with Dr. Felix Feng (Month 12-18).

Progress (100% completed)

We performed *in-vivo* tumor formation assays with PCAT29-KD, SCHLAP-KD and PCAT51 KD cell lines. Results for PCAT29 and SCHLAP are now published (see appendix). As shown below PCAT51 KD using shRNA reduced tumor growth *in-vivo*.

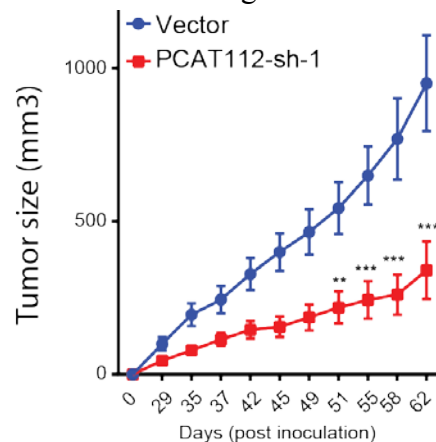


Figure 6: Growth of LNCaP-AR xenografts stably expressing control shRNA or shRNA targeting PCAT51. Data, mean \pm SEM. *, $P < 0.05$ by the one way ANOVA.

Task 9: To explore the role of androgen signaling in regulation of lncRNAs (Month 12-18)

To address the role of AR in lncRNA regulation, we performed an unbiased discovery of androgen receptor regulated long non-coding RNA. Transcriptome sequencing (RNA-seq) on DHT stimulated AR dependent cell lines VCaP and LNCaP was performed. The figure below describes the outline of the experiment performed.

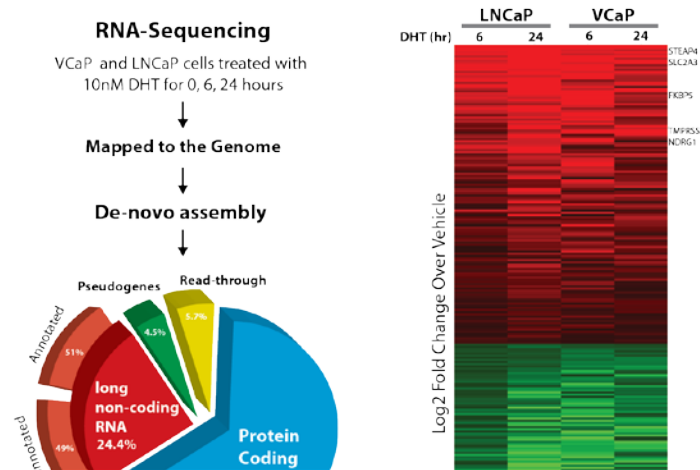


Figure 7: Schematic representation of experimental outline. B. Heatmap of gene expression changes after DHT stimulation in LNCaP and VCaP cells. C. Validation of 50 coding and non-coding transcripts by q-RT-PCR.

Specific Aim 2: To elucidate the function of lncRNA in metastatic castrate resistant prostate cancer: (Month 6-24)

Task 1: Microarray analysis (Month 6-12)

1. Perform microarray after lncRNA knockdown using siRNA.
2. Training with bioinformaticians on analysis of microarray data

Progress (100% completed)

We performed Microarray after KD of PCAT29 (Figure-8) and PCAT51 (Figure -9). The data was analyzed and Gene-Set enrichment analysis was performed to find relevant concepts.

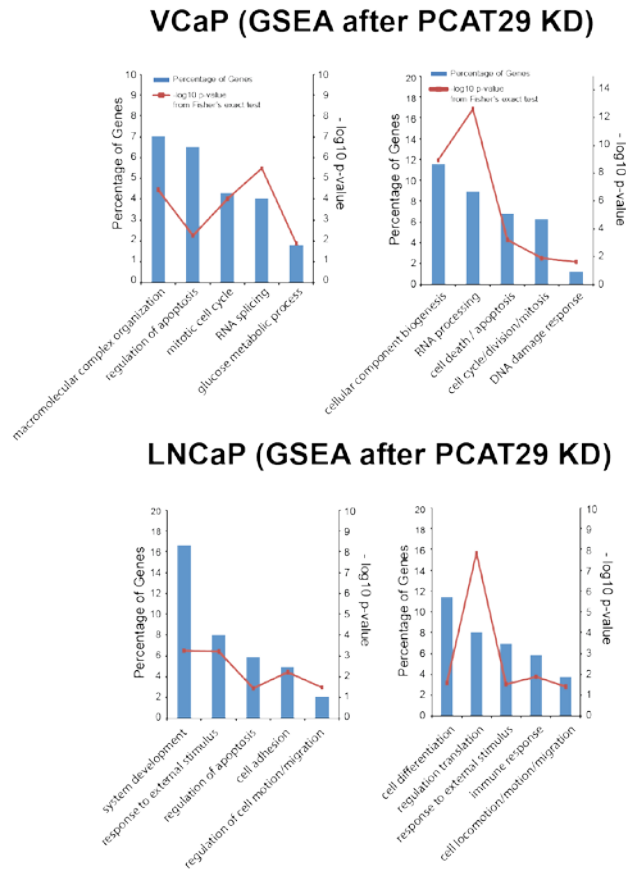


Figure 8: GSEA analysis of Micorarray data obtained after PCAT29 knockdown in VCaP and LNCaP cells. Bars represent the Geo categories enriched.
PCAT51 KD in MDA_PCa_2B cells

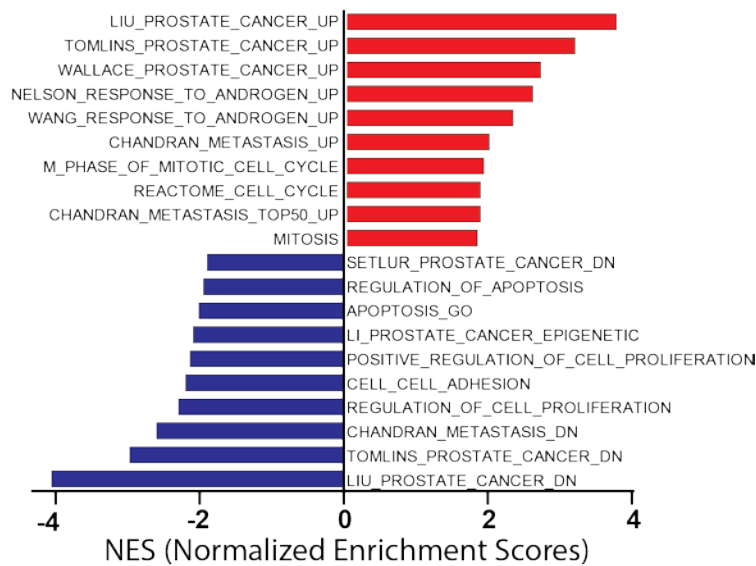


Figure 9: GSEA analysis of Micorarray data obtained after PCAT51 knockdown in MDA_PCa_2B cells. Bars represent the Geo categories enriched.

Training: I did training with bioinformatician in lab and learned how to analyze microarray data. Learned to use LIMMA package in R. Other software's such as MeV4 and Treeview were also learned for heat map generation. Further, I am taking courses on "R" with University of Michigan biostatistics department.

Task 2: To make tilling probes for lncRNAs. (Month 9-12)

Progress (100% completed)

We have published ChIRP technique with SCHLAP (see appendix). For PCAT51 we made 10 tiling arrays. By performing ChIRP we have determined that probe1 and 2 are the best. These probes will be use in subsequent experiments.

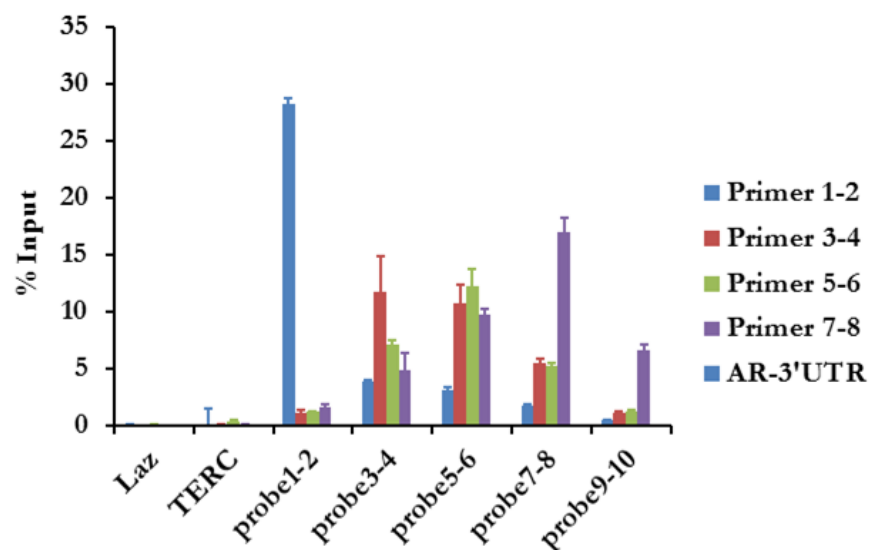


Figure 10: PCAT51 was pull-down using 5 set of 2 probes each. Enrichment of PCAT51 compared to input was determined by q-RT-PCR

Task 3: To standardize ChiRP protocol. (Month 9-12)

Progress (100% completed)

ChIRP protocol has now been standardized. We published this in SCHLAP manuscript (see appendix)

Task 4: To perform ChiRP and Mass spectrometry analysis (Month 12-18).

Progress (100% completed)

After few rounds of optimization we have decided to approach this problem using a different technique. In order to identify RNA interacting proteins we are now performing in-vitro RNA pull-downs following mass-spectrometry. We used PCAT51 as our candidate for this approach. Briefly, RNA was synthesized in-vitro and labeled with BrU. Labelled RNA was then used to pulldown protein. Mass-spectrometry was used to identify the proteins bound to RNA.

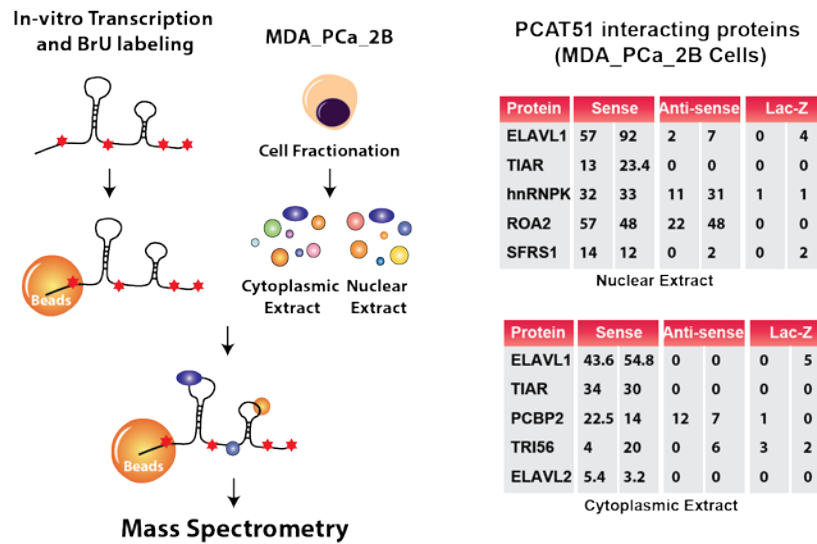
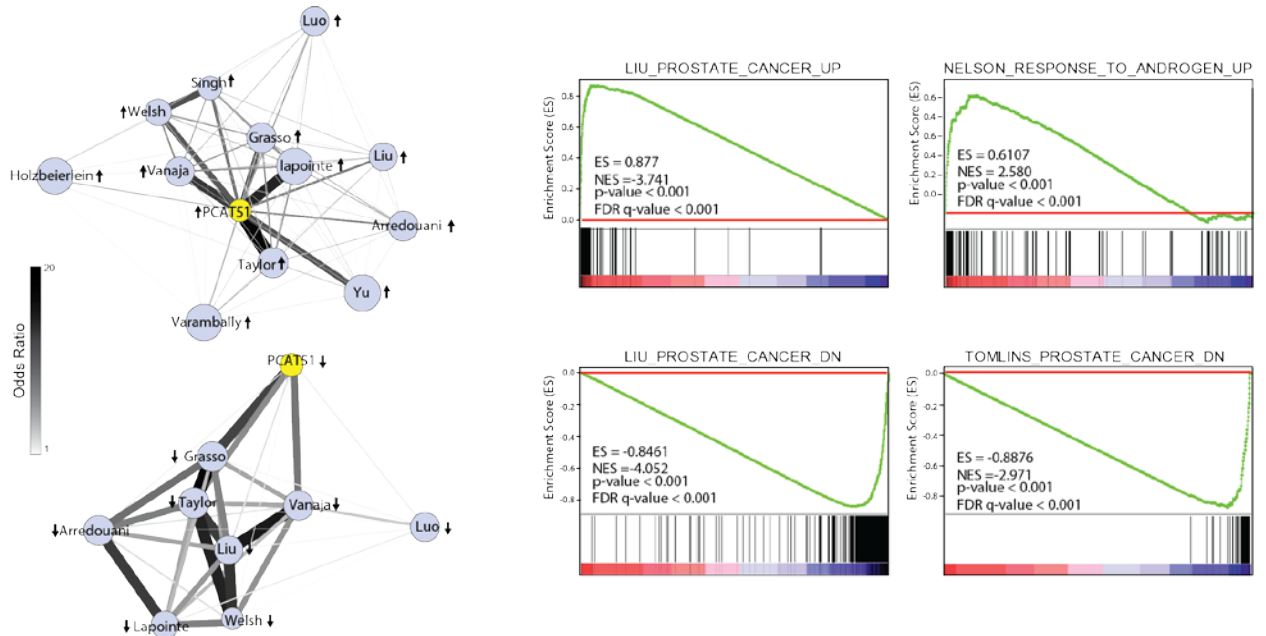


Figure 11: Schematic representation of the BrU pulldown approach. List of protein in nucleus and cytoplasm that were found to be associate with PCAT51.

Task 3: Correlation analysis (Month 12-18): Correlation data for PCAT29 and SCHLAP are published. A similar analysis was performed for PCAT51. Correlated genes were then analysed by GSEA and oncomine. As shown below positively correlated genes were significantly enriched in prostate cancer concepts suggesting role of PCAT51 in cancer progression.



Task 4: Validation of microarray and ChIP results (Month 18-24).

Progress (0% completed)

In progress

Specific Aim 3: To identify lncRNAs that serve as potential biomarkers of disease progression. (Month 12-24)

Task 1: Screening of lncRNA in prostate cancer cohorts (Month 12-24): We have screened PCAT29 and SCHLAP in cohort generated using samples from University of Michigan. Other lncRNA candidates will be screened in this cohort. Other independent prostate cancer cohorts will also be screened for lncRNA expression.

Progress (100% completed)

Expression of PCAT51 was assessed in RNA-seq data obtained from University of Michigan as well as other sources such as TCGA and other prostate cancer studies. As shown below PCAT51 was significantly expressed in prostate cancer compared to benign samples.

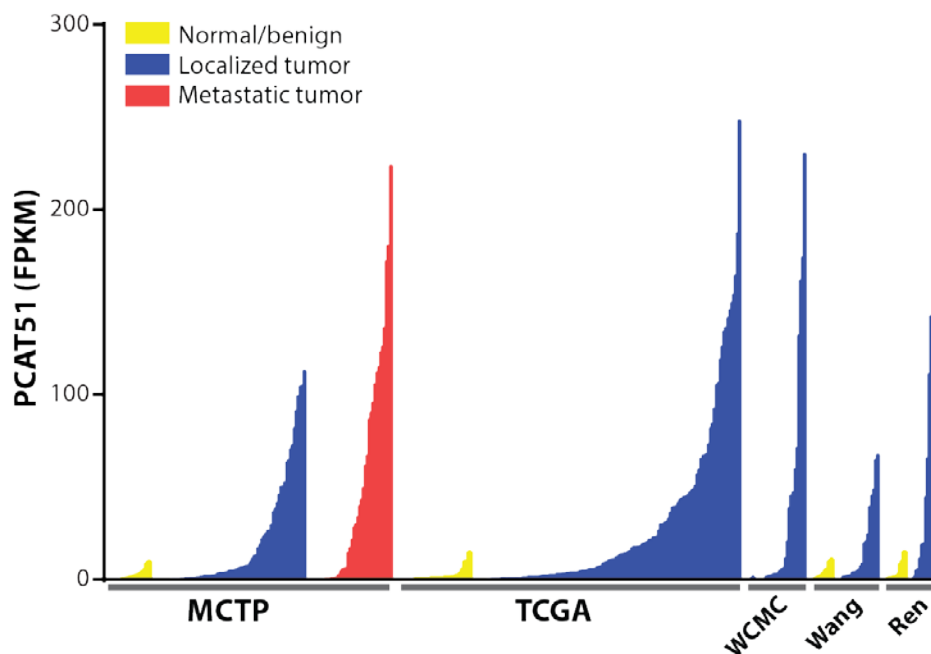


Figure 11: Expression of PCAT51 in various prostate datasets.

Task 2: Screening of lncRNAs in FFPE cohort (Month 12-24).

Progress (0% completed)

Task 3: Fresh Frozen prostate tissue cohort: Samples from Rapid Autopsy program will be obtained in collaboration with Dr. Pienta and lncRNA expression will be measured followed by analysis similar to FFPE cohort. (Month 1-24).

Progress (0% completed)

Task 4: Microarray analysis using data from Genome DX. (Month 18-24)

Progress (0% completed)

What opportunities for training and professional development has the project provided?

Several opportunities for training and professional development were provided while working on this project.

One-on-one meeting with mentor: Data was discussed with mentor (Dr. Arul Chinnaiyan) in biweekly individual meetings and via monthly progress reports. Constant inputs and advice were provided.

Presentations at Scientific Conference: Data obtained was presented at several scientific conferences:

1. Annual AACR conference.
2. Keystone Symposia Conference. “Long Noncoding RNAs: Marching toward Mechanism” 2014
3. Sixth Annual Prostate Cancer Program Retreat

Manuscript writing: A manuscript describing the results was published earlier this year.

Weekly lab meetings: All important research findings were presented during weekly lab meetings in front of the group.

How were the results disseminated to communities of interest?

Results from the grant were published in form on a manuscript as well as in form of conference proceeding. (see appendix)

What do you plan to do during the next reporting period to accomplish the goals?

By next reporting period I anticipate to accomplish all the aims proposed in the study. We also anticipate to have published another manuscript describing results on lncRNA PCAT51.

3. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

Once completed, the proposed project “**Biological and clinical translation of novel long non-coding RNAs (lncRNAs) associated with metastatic castration-resistant prostate cancer**” will have substantial impact on understanding prostate cancer biology and its clinical management. We have proposed to study the role of long non-coding RNAs (lncRNAs) in development and progression of prostate cancer as well as explore the potential of lncRNA to serve as biomarkers to predict disease progression.

Till data we have discovered several new lncRNA molecules that play important role in prostate cancer progression. One such lncRNA, *PCAT29*, was shown to be the first androgen receptor–

repressed lncRNA that functions as a tumor suppressor. We also show that loss of PCAT29 may identify a subset of patients at higher risk for disease recurrence. PCAT29 based biomarkers can be useful in prostate cancer management in the clinics.

We anticipate that at the end of this proposed project, we will be able to discover several such lncRNAs that can predict disease progression and will be able to stratify patient into high risk vs low risk disease.

What was the impact on other disciplines?

Nothing to report

What was the impact on technology transfer?

Nothing to report

What was the impact on society beyond science and technology?

Nothing to report

4. CHANGES/PROBLEMS::

Changes in approach and reasons for change

To identify protein that binds to lncRNA we initially proposed to perform ChIRP following mass spectrometry. However, due to incompatibility of ChIRP reagents with mass-spec, we utilized a slightly different approach. We performed pull-down using in-vitro labeled RNA. Interacting proteins were then identified by mass-spectrometry.

Actual or anticipated problems or delays and actions or plans to resolve them

None

Changes that had a significant impact on expenditures

None

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

No Changes

Institutional Animal Care and Use Approval Number and Dates:

PRO00004337, 12/12/2012 – 12/12/2015 (Mouse)

Significant changes in use or care of human subjects:

None

Significant changes in use or care of vertebrate animals.

None

Significant changes in use of biohazards and/or select agents.

None

5. PRODUCTS:

Publications, conference papers, and presentations

Journal publications:

Malik R, Patel L, Prensner JR, Shi Y, Iyer M, Subramaniyan S, Carley A, Niknafs YS, Sahu A, Han S, Ma T, Liu M, Asangani IA, Jing X, Cao X, Dhaneshekaran SM, Robinson D, Feng FY, Chinnaiyan AM. The lncRNA PCAT29 Inhibits Oncogenic Phenotypes in Prostate Cancer. *Mol Cancer Res*. 2014 Aug;12(8):1081-7. doi: 10.1158/1541-7786.MCR-14-0257. PMCID: PMC4135019.

Acknowledgement of federal support: Yes.

Prensner, J.R.*, A. Sahu*, M.K. Iyer*, **Malik R.***, B. Chandler, I.A. Asangani, A. Poliakov, I.A. Vergara, M. Alshalalfa, R.B. Jenkins, E. Davicioni, F.Y. Feng, and A.M. Chinnaiyan. 2014b. The lncRNAs PCGEM1 and PRNCR1 are not implicated in castration resistant prostate cancer. *Oncotarget*. (*Co-first Author)

Acknowledgement of federal support: Yes.

Other publications, conference papers, and presentations.

Rohit Malik, Matthew K. Iyer, John R. Prensner, Lalit Patel, Sumin Han, Wei Chen, Felix Feng, Arul M. Chinnaiyan. Identification and characterization of a novel androgen-regulated long non-coding RNA in prostate cancer. [abstract]. In: Proceedings of the 104th Annual Meeting of the American Association for Cancer Research; 2013 Apr 6-10; Washington, DC. Philadelphia (PA): AACR; Cancer Res 2013;73(8 Suppl):Abstract nr 1120. doi:10.1158/1538-7445.AM2013-1120 (*)

Rohit Malik, Matthew K Iyer, Shruthi Subramaniam, Yasuyuki Hosono, Anirban Sahu, Xia Jiang, Yang Shi, Vishal Kothari, Xuhong Cao, Dan Robinson, Saravana M. Dhanasekaran, Felix Y Feng and Arul M Chinnaiyan. Keystone Symposia Conference. “Long Noncoding RNAs: Marching toward Mechanism” 2014

Rohit Malik, Matthew K. Iyer, John R. Prensner, Lalit Patel, Sumin Han, Wei Chen, Felix Feng, Arul M. Chinnaiyan. “Sixth Annual Prostate Cancer Program Retreat”. March 18-20, 2013(*)

Technologies or techniques

Nothing to Report.

Inventions, patent applications, and/or licenses

Nothing to Report.

Other Products

Nothing to Report.

6. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name:	Rohit Malik – No change
-------	-------------------------

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

New Award:

Award No: N/A (PI: Malik) 01/20/2014 – 01/20/2017 0 cal months effort

Sponsor: Prostate Cancer Foundation Young Investigator Award

Title: Characterization and therapeutic targeting of androgen receptor co-activators in castration resistance prostate cancer

Goals: The goals of this project are three fold. 1) To develop highly potent small molecule inhibitors of the menin-MLL interaction with optimized drug-like properties. 2) To study the mechanistic effects of pharmacologic menin-MLL inhibition on AR regulated signaling and tumor growth. 3) To establish the efficacy of menin-MLL inhibitors *in vivo* employing CRPC pre-clinical models.

Overlap: None

Role: PI, Young Investigator

7. **SPECIAL REPORTING REQUIREMENTS**

- **None**

8. **APPENDICES:** *Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.*

1. Malik et. al., MCR 2014
2. Prensner et. al., Nat Gen 2013
3. CV_Rohit Malik

Molecular Cancer Research



The lncRNA *PCAT29* Inhibits Oncogenic Phenotypes in Prostate Cancer

Rohit Malik, Lalit Patel, John R. Prensner, et al.

Mol Cancer Res 2014;12:1081-1087. Published OnlineFirst July 16, 2014.

Updated version	Access the most recent version of this article at: doi: 10.1158/1541-7786.MCR-14-0257
Supplementary Material	Access the most recent supplemental material at: http://mcr.aacrjournals.org/content/suppl/2014/07/19/1541-7786.MCR-14-0257.DC1.html

Visual Overview	A diagrammatic summary of the major findings and biological implications: http://mcr.aacrjournals.org/content/12/8/1081/F1.expansion.html
------------------------	---

Cited Articles	This article cites by 15 articles, 4 of which you can access for free at: http://mcr.aacrjournals.org/content/12/8/1081.full.html#ref-list-1
-----------------------	--

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org .

The lncRNA *PCAT29* Inhibits Oncogenic Phenotypes in Prostate Cancer

Rohit Malik^{1,2}, Lalit Patel^{1,2}, John R. Prensner^{1,2}, Yang Shi^{1,3}, Matthew K. Iyer^{1,2}, Shruthi Subramaniyan¹, Alexander Carley¹, Yashar S. Niknafs^{1,2}, Anirban Sahu^{1,2}, Sumin Han^{1,4}, Teng Ma^{1,4}, Meilan Liu⁴, Irfan A. Asangani^{1,2}, Xiaojun Jing^{1,2}, Xuhong Cao^{1,2}, Saravana M. Dhanasekaran^{1,2}, Dan R. Robinson^{1,2}, Felix Y. Feng^{1,4,5}, and Arul M. Chinnaiyan^{1,2,5,6}

Abstract

Long noncoding RNAs (lncRNA) have recently been associated with the development and progression of a variety of human cancers. However, to date, the interplay between known oncogenic or tumor-suppressive events and lncRNAs has not been well described. Here, the novel lncRNA, prostate cancer–associated transcript 29 (*PCAT29*), is characterized along with its relationship to the androgen receptor. *PCAT29* is suppressed by DHT and upregulated upon castration therapy in a prostate cancer xenograft model. *PCAT29* knockdown significantly increased proliferation and migration of prostate cancer cells, whereas *PCAT29* overexpression conferred the opposite effect and suppressed growth and metastases of prostate tumors in chick chorioallantoic membrane assays. Finally, in prostate cancer patient specimens, low *PCAT29* expression correlated with poor prognostic outcomes. Taken together, these data expose *PCAT29* as an androgen-regulated tumor suppressor in prostate cancer.

Implications: This study identifies *PCAT29* as the first androgen receptor–repressed lncRNA that functions as a tumor suppressor and that its loss may identify a subset of patients at higher risk for disease recurrence.

Visual Overview: <http://mcr.aacrjournals.org/content/early/2014/07/31/1541-7786.MCR-14-0257/F1.large.jpg>.

Mol Cancer Res; 12(8); 1081–7. ©2014 AACR.

Introduction

Recently, data from the ENCODE project have revealed that the majority of the transcriptome is composed of noncoding RNAs (1). While the classification of these noncoding RNAs is still in development, long noncoding RNAs (lncRNA) are of particular interest, given the similar features they share with protein-coding genes as well as recent evidence of their roles in cancer biology (2, 3). lncRNAs are polyadenylated RNA species that are more

than 200 bp in length, transcribed by RNA polymerase II, and associated with common epigenetic signatures such as of histone 3 lysine 4 trimethylation (H3K4me3) at the transcriptional start site (TSS) and histone 3 lysine 36 trimethylation (H3K36me3) in the gene body (4). Several lncRNAs have been shown to play a role in biologic processes such as X-chromosomal inactivation, pluripotency (5), and gene regulation (6). Recently, several lncRNAs have been implicated in cancer initiation and progression (3, 7). Apart from their role in tumor initiation and progression, lncRNAs have been shown to be promising biomarkers. In prostate cancer, PCA3 is a well-studied prostate cancer biomarker that is now available for clinical use as a urine biomarker assay for diagnosis of prostate cancer (8, 9).

Despite their involvement in various cellular processes, the majority of lncRNAs are uncharacterized, and their role in cancer initiation and progression remains unclear. Using transcriptome sequencing, our group recently identified more than 100 lncRNAs, named prostate cancer–associated transcripts (*PCATs*), which are differentially expressed or have outlier profiles in prostate cancer versus normal tissue (3). Here we find that one of these novel lncRNAs, *PCAT29*, exhibits cancer-suppressive phenotypes, including inhibition of cell proliferation, migration, tumor growth, and metastases. In accordance with this, *PCAT29* is repressed by androgen signaling, and low *PCAT29* expression associates with worse clinical outcomes.

¹Michigan Center for Translational Pathology, University of Michigan, Ann Arbor, Michigan. ²Department of Pathology, University of Michigan, Ann Arbor, Michigan. ³Department of Biostatistics, University of Michigan, Ann Arbor, Michigan. ⁴Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan. ⁵Comprehensive Cancer Center, University of Michigan, Ann Arbor, Michigan. ⁶Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan.

Note: Supplementary data for this article are available at Molecular Cancer Research Online (<http://mcr.aacrjournals.org/>).

F.Y. Feng and A.M. Chinnaiyan share senior authorship of this article.

Corresponding Authors: Arul M. Chinnaiyan, Comprehensive Cancer Center, University of Michigan Medical School, 1400 E. Medical Center Dr. 5316 CCGC 5940, Ann Arbor, MI 48109. Phone: 734-615-4062; Fax: 734-615-4055; E-mail: arul@med.umich.edu; and Felix Y. Feng, Department of Radiation Oncology, University of Michigan Medical Center, 1500 East Medical Center Drive, UHB2C490-SPC5010, Ann Arbor, MI 48109. Phone: 734-936-4302; Fax: 734-763-7371; E-mail: ffeng@med.umich.edu

doi: 10.1158/1541-7786.MCR-14-0257

©2014 American Association for Cancer Research.

Materials and Methods

Cell lines and reagents

Prostate cancer cells were cultured as follows: VCaP cells in DMEM with GlutaMAX (Invitrogen) and LNCaP and DU145 cells in RPMI-1640 (Invitrogen) in a 5% CO₂ cell culture incubator. All the media were supplemented with 10% FBS (Invitrogen) and 1% penicillin–streptomycin (Invitrogen). All cell lines were purchased from ATCC and were authenticated.

For stable knockdown of *PCAT29*, LNCaP and VCaP cells were transfected with lentiviral constructs encoding 2 different *PCAT29* shRNAs or nontargeting shRNAs in the presence of polybrene (8 µg/mL; Supplementary Table S1A). After 48 hours, transduced cells were grown in culture media containing 3 to 5 µg/mL puromycin. For *PCAT29* overexpression, 2 isoforms of *PCAT29* were generated by subcloning the PCR product into the *CPOI* sites of the pCDH-CMV vector (System Biosciences). Five hundred base pairs of the genomic region was attached at the 5' end of each isoform. Lentiviral particles were made and DU145 cells were transduced as described above.

Gene expression by quantitative PCR

Total RNA was isolated using TRIzol (Invitrogen) and an RNeasy kit (Qiagen) according to manufacturers' instruction. Total RNA was reverse transcribed into cDNA using SuperScript III and random primers (Invitrogen). Quantitative PCR (qPCR) was performed using SYBR Green Master Mix (Applied Biosystems) on an Applied Biosystems 7900HT Real-Time System. The relative quantity of the target gene was computed for each sample using the $\Delta\Delta C_t$ method by comparing mean C_t of the gene to the mean C_t of the housekeeping gene *GAPDH*. All the primers were obtained from Integrated DNA Technologies (IDT). Sequences of all the primers used are listed in Supplementary Table S1B.

Rapid amplification of cDNA ends

5' and 3' RACE was performed using the GeneRacer RLM-RACE kit (Invitrogen) following manufacturer's instruction. RACE PCR products were separated on a 1% agarose gel. Individual bands were gel purified, cloned in pcr4-TOPO vector, and sequenced using M13 primers.

Expression of *PCAT29* after castration in prostate tumor xenograft model

Five-week-old male nude athymic BALB/c *nu/nu* mice (Charles River Laboratory) were used for xenograft studies. LNCaP cells were resuspended in 100 µL of PBS with 20% Matrigel (BD Biosciences) and implanted subcutaneously into the left flank regions of the mice. Mice were castrated and euthanized 5 days after castration. RNA was extracted from the xenografts and expression of *PCAT29* and *FKBP5* was measured. All experimental procedures involving mice were approved by the University Committee on Use and Care of Animals at the University of Michigan (Ann Arbor, MI) and conform to their relevant regulatory standards.

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) was performed with polyclonal androgen receptor antibody (Millipore PG21) using HiCell ChIP kit (Diagenode) following manufacturer's instruction. Briefly, cells were treated with 10 µmol/L MDV3100 or 10 µmol/L bicalutamide 16 hours before the treatment with 10 nmol/L DHT for 12 hours. Approximately 1 million cells were cross-linked per antibody with 1% formaldehyde. Chromatin was sonicated to an average length of 500 bp and centrifuged to remove debris. Magnetic protein-G beads were coated with 6 µg of antibody and incubated with chromatin overnight at 4°C. Protein–chromatin–antibody complexes were washed thrice and cross-linking was reversed. ChIP products were cleaned using IPure kit (Diagenode). Eluted DNA was quantified by RT-PCR using primers described in Supplementary Table S1B.

Cell proliferation and migration assay

LNCaP and DU145 cells stably expressing *PCAT29* shRNA-1 and 2 or *PCAT29* isoform 1 and 2 were seeded in 24-well plates. Cells were trypsinized and counted by using Coulter Counter (Beckman Coulter) at the indicated time points in triplicate. For migration assays, approximately 1×10^5 cells were seeded in the upper chamber of a Boyden chamber. About 500 µL of complete medium (10% FBS) was added to the lower chamber as a chemoattractant. Forty-eight hours after seeding, cells on the upper surface were removed using a cotton swab. Inserts were fixed with 3.7% formaldehyde and migrated cells on the lower surface of the membrane were stained with crystal violet. The inserts were treated with 10% acetic acid, and absorbance was measured at 560 nm.

Gene expression microarray

Expression profiling of VCaP and LNCaP cells after *PCAT29* knockdown was performed using the Agilent Whole Human Genome Oligo Microarray as described (7). GEO accession number: GSE58397.

Chicken chorioallantoic membrane assay

22RV1 cells were transduced with empty vector (pcDH) or *PCAT29*-isoform-1. A total of 10^6 cells were inoculated on the chicken chorioallantoic membrane (CAM) assay as described previously (10). For tumor growth and metastasis, the eggs were incubated for 18 days in total, after which the extra-embryonic tumor were exercised and weighed, and the embryonic livers were harvested and analyzed for the presence of tumor cells by quantitative human Alu-specific PCR. Quantification of human cells in the extracted DNA was performed as described (11). Fluorogenic TaqMan qPCR probes were applied as above and DNA copy numbers were quantified.

Kaplan–Meier analysis of *PCAT29*

For outcomes analysis, *PCAT29* expression was determined on a cohort of 51 radical prostatectomy specimens from patients with prostate cancer at the University of

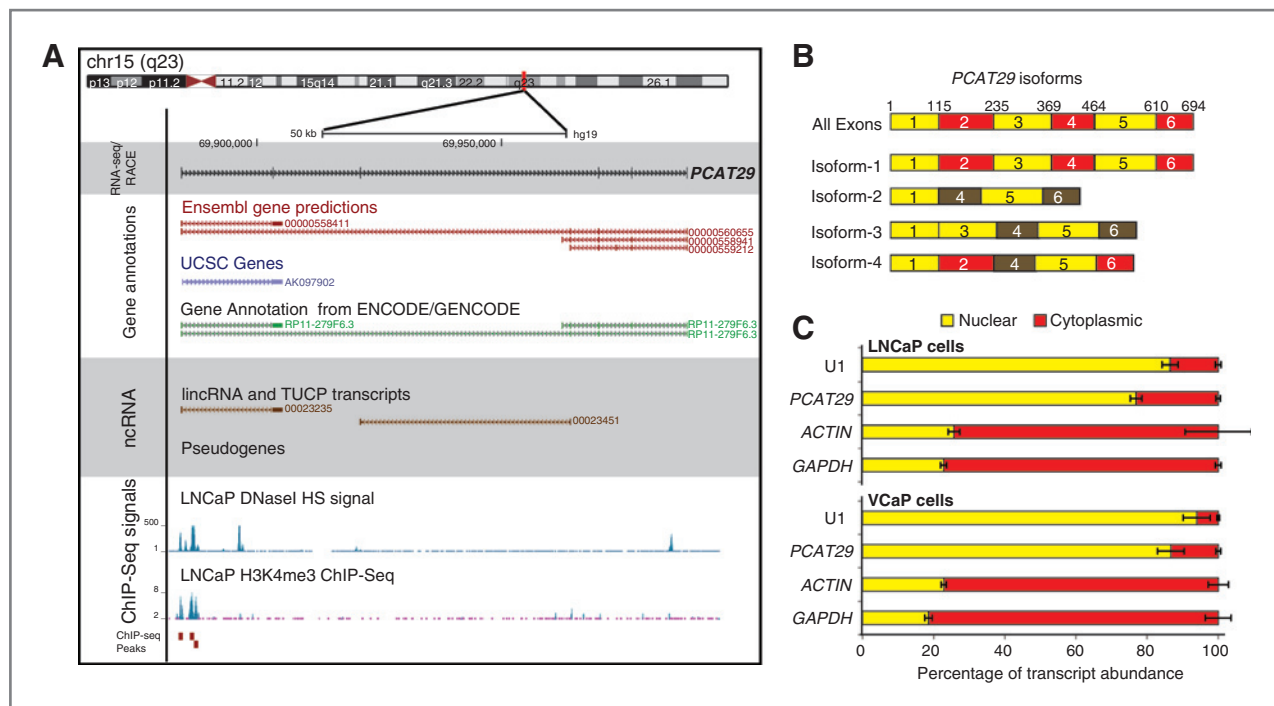


Figure 1. Characterization of *PCAT29*. A, genome browser representation of *PCAT29*. Current gene annotations from Ensembl, ENCODE, UCSC genes and lncRNA databases are shown. ChIP-Seq data for H3K4me3 and DNaseI HS signal in LNCaP cells obtained from ENCODE. B, schematic representation of *PCAT29* isoforms as determined by RACE analysis. C, nuclear and cytoplasmic distribution of various noncoding and protein-coding transcripts in LNCaP and VCaP cells. Error bar, \pm SEM.

Michigan with long-term biochemical recurrence outcomes. Biochemical recurrence was defined by an increase of PSA of 0.2 ng/mL over the PSA nadir following prostatectomy. *PCAT29* expression was determined by a SYBR-Green qPCR assay using the average of *GAPDH* + *HMBS* for data normalization using the $\Delta\Delta C_t$ method. Expression data were transformed using a z -score and patients were defined as high (top 33% of patients) or low (bottom 66% of patients) for *PCAT29* expression. Kaplan–Meier curves for biochemical recurrence-free survival were generated for *PCAT29*-high and *PCAT29*-low patients using the GraphPad Prism program. Statistical significance was determined with a log-rank test.

Results

PCAT29 is a novel long nuclear noncoding RNA

Using RNA-Seq data from prostate cancer tissues, we previously identified 121 lncRNAs, named PCATs, which demonstrate differential expression or outlier profiles in prostate cancer compared with normal tissue (3). Here we characterize and functionally investigate one of the top outlier lncRNAs, *PCAT29* (Ensembl ID ENSG00000259641). Using the predicted transcript structures, we designed exon spanning primers and performed rapid amplification of cDNA ends (RACE) to determine the full exon structure. As shown in a genome browser view, *PCAT29* is a 694-bp polyadenylated transcript present on chr15(q23), and the *PCAT29* gene spans over a 10-kb stretch (Fig. 1A; Supple-

mentary Fig. S1A). *PCAT29* is composed of 6 exons that are alternatively spliced to produce multiple isoforms (Fig. 1B). To further characterize *PCAT29*, we interrogated recently published ENCODE data for H3K4 trimethylation (H3K4me3) and DNaseI hypersensitive sites (DNaseH), marks that predicts for open chromatin state and are commonly found near or at the TSSs, generated in the prostate cancer cell line LNCaP (4). We found several DNaseH and H3K4 trimethylation peaks at the TSS of *PCAT29*, suggesting that *PCAT29* is an actively transcribed gene (Fig. 1A).

To confirm that *PCAT29* is indeed a noncoding RNA, we assessed the protein-coding potential of *PCAT29* using the coding potential calculator (CPC) algorithm, which discriminates coding genes (positive score) from noncoding transcripts (negative score; ref. 12). *PCAT29* had a CPC score of -0.8921 , whereas protein-coding genes such as *TP53* and β -actin scored $+8.25$ and $+3.70$, respectively (Supplementary Fig. S1B). Consistent with this finding, we found that in both LNCaP and VCaP cells, expression of *PCAT29* was limited to nucleus, whereas other protein-coding mRNAs, such as *GAPDH* and β -actin, were expressed in cytoplasm (Fig. 1C). We then verified the expression of *PCAT29* in various prostate cancer cell lines (LNCaP, VCaP, 22RV1, DU145, PC3) and immortalized or primary prostate epithelial cells (RPWE and PrEC). *PCAT29* expression was highest in androgen receptor-dependent cell lines such as LNCaP, VCaP, and 22RV1 (Supplementary Fig. S1C).

Next, we assessed the expression of *PCAT29* in various tissues using transcriptome sequencing data. *PCAT29* expression, although not limited to prostate, was enriched in prostate samples compared with other tissues (Supplementary Fig. S1D).

Androgen receptor binds to the *PCAT29* promoter and regulates *PCAT29* expression

We next examined the effect of androgen receptor signaling on *PCAT29* in LNCaP cells stimulated with 10 nmol/L DHT. As shown in Fig. 2A, *PCAT29* expression was suppressed upon stimulation with DHT in a time-dependent fashion both in LNCaP and VCaP cells. In contrast, expression of canonical androgen receptor target genes, such as *FKBP5* and *KLK3*, was increased upon stimulation (Supplementary Fig. S2A). To examine whether the suppression of *PCAT29* was androgen receptor-specific, LNCaP cells

were pretreated with the androgen receptor antagonists MDV3100 or bicalutamide before treatment with DHT. As expected, DHT stimulation suppressed the expression of *PCAT29*, and pretreatment with MDV3100 or bicalutamide rescued this suppression. Similarly, expression of *PCAT29* in LNCaP cells grown in charcoal-stripped media as well as in an androgen receptor-independent variant of LNCaP cells (C42) was higher than in cells grown in serum-containing media and LNCaP cells, respectively (Supplementary Fig. S2B and S2C). We next investigated whether androgen receptor suppresses the expression of *PCAT29* *in vivo*. LNCaP xenografts were established in mice followed by physical castration to ablate androgen receptor signaling. As expected, 5 days of castration led to significant increase in the expression of *PCAT29* in tumors (Fig. 2C). In contrast, expression of *FKBP5* was reduced in tumors from castrated mice. Taken together, our results suggest that stimulation

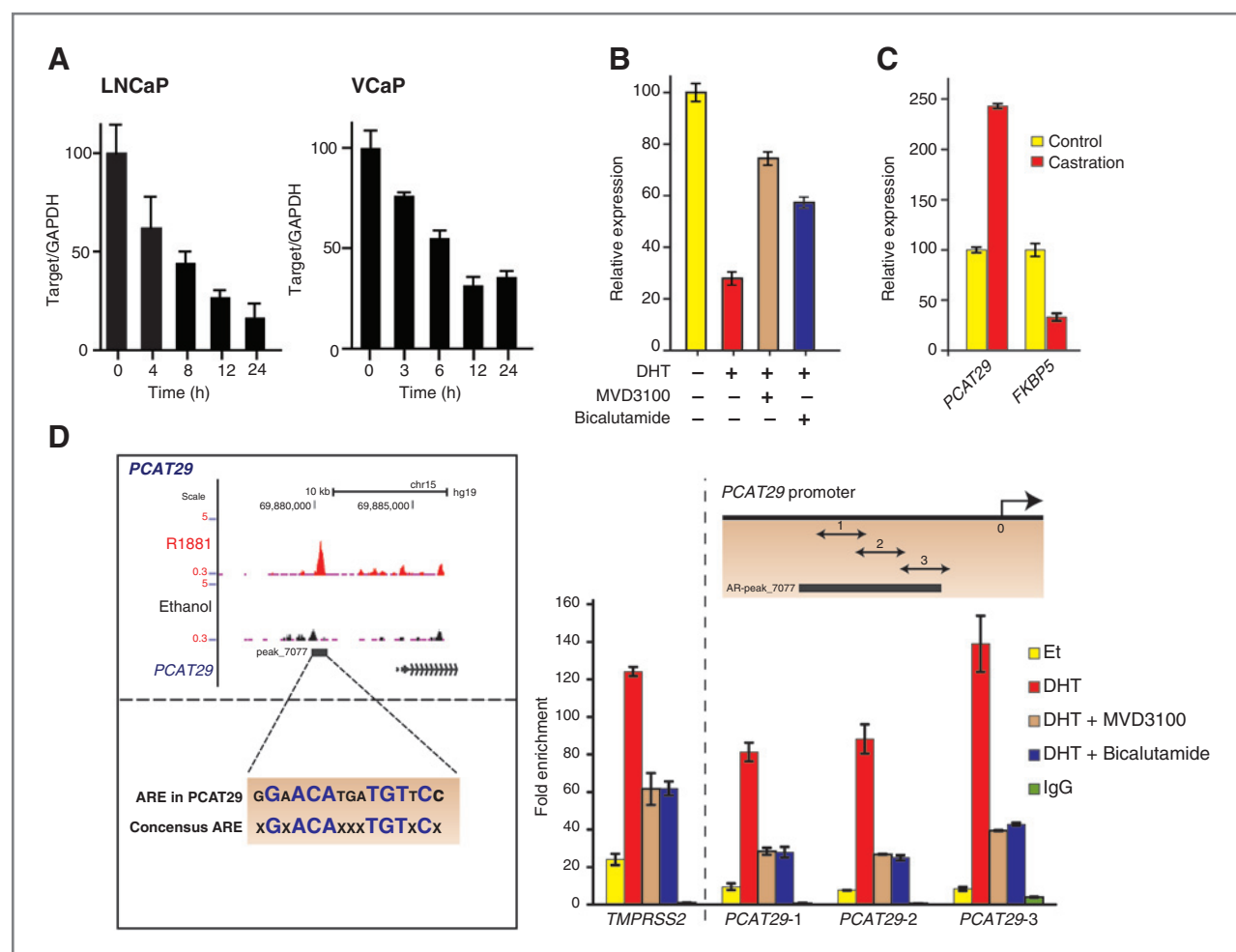


Figure 2. Androgen receptor binds to the promoter of *PCAT29* and suppresses its expression. A, expression of *PCAT29* in LNCaP and VCaP cells treated with 10 nmol/L DHT for indicated time points. B, expression of *PCAT29* in LNCaP cells treated with 10 nmol/L DHT in the presence or absence of 10 μ mol/L MDV3100 or bicalutamide. C, expression of *PCAT29* and *FKBP5* in LNCaP xenografts obtained from control mice and mice that were physically castrated for 5 days. D, genome browser representation of androgen receptor (AR) binding on the promoter of *PCAT29* before and after stimulation with 1 nmol/L R1881. Consensus androgen-responsive elements (ARE) and ARE present in the *PCAT29* promoter are shown. Inset, ChIP-PCR to confirm AR occupancy on *TMPRSS2* and *PCAT29* gene promoter. The y-axis represents AR ChIP enrichment in VCaP cells treated with 10 nmol/L DHT normalized to ethanol (Et)-treated cells. Bars, SEM.

of androgen receptor leads to suppression of *PCAT29* expression.

To further study the association of *PCAT29* expression with androgen signaling, we interrogated published ChIP-Seq data (13) and found androgen receptor-binding sites in the promoter region of *PCAT29* (Fig. 2D). These peaks were similar to those observed in other known androgen receptor-regulated genes (Supplementary Fig. S2D). Upon closer inspection, we found a canonical androgen receptor-binding site near the *PCAT29* TSS in a putative enhancer region bound by androgen receptor (Fig. 2D). We confirmed our ChIP-Seq data by performing ChIP for androgen receptor followed by PCR for the *PCAT29* promoter. As shown in Fig. 2D, stimulation of VCaP cells with DHT led to an increase in association of androgen receptor with the *PCAT29* promoter. This association was reduced in cells pretreated with bicalutamide and MDV3100. Taken together, our data suggest that androgen receptor can directly bind to the promoter of *PCAT29* and leads to the suppression of gene expression.

PCAT29* regulates oncogenic phenotypes *in vitro* and *in vivo

The androgen receptor drives oncogenesis in treatment-naïve prostate cancer as well as disease progression in castration-resistant prostate cancers. Because androgen receptor binds to the *PCAT29* promoter and regulates gene expression, we investigated the functional role of *PCAT29*. Two independent shRNAs were designed to knockdown the expression of *PCAT29* in cells (Supplementary Fig. S3A and S3B). VCaP and LNCaP cells were transfected with *PCAT29* shRNAs following analysis using gene expression microarray. We found GO concepts enriched for cell cycle, proliferation, and migration-related genes, suggesting a role of *PCAT29* in proliferation and migration (Supplementary Fig. S3D–S3G). Next, we defined a signature of genes positively and negatively correlated with *PCAT29* expression from prostate cancer samples as described before (7). We checked the overlap of these genes with the top 1500 differentially expressed genes in *PCAT29* knockdown samples of VCaP and LNCaP cells. As expected, the positively correlated genes show a significant overlap with genes downregulated with knockdown of *PCAT29* and the negatively correlated genes show a significant overlap with genes upregulated by knockdown of *PCAT29* in both VCaP and LNCaP ($P < 0.001$ for all pairwise comparisons of overlapping genes, Supplementary Fig. S4A–S4D). For overlapping genes, we did see enrichment in pathways such as cell cycle, apoptosis, and cell growth (Supplementary Fig. S4A–S4D). Taken together, this analysis suggested a role of *PCAT29* in cell proliferation and migration.

To experimentally validate this observation, cell proliferation was assessed in LNCaP cells transfected with control versus *PCAT29* shRNAs. To our surprise, knockdown of *PCAT29* in LNCaP cells led to an increase in cell proliferation and migration (Fig. 3A). To further validate this observation, we stably overexpressed the 2 most prevalent isoforms of *PCAT29* in DU145 prostate cancer cells using a

lentiviral vector (Supplementary Fig. S3C). Consistent with the previous knockdown studies, overexpression of these 2 isoforms of *PCAT29* in DU145 led to suppression of cell proliferation and migration (Fig. 3B). We next assessed whether similar effects of *PCAT29* could be achieved *in vivo*. 22RV1 prostate cancer cells overexpressing *PCAT29* (isoform-1) were implanted on the CAM of a chicken egg. Compared with control cells, overexpression of *PCAT29* significantly decreased the growth of tumor on the CAM as well as decreased liver metastases (Fig. 3C).

Finally, we measured the expression of *PCAT29* in an independent cohort of 51 radical prostatectomy specimens from patients with prostate cancer with localized disease and clinical follow-up. As shown in Kaplan–Meier analysis (Fig. 3D), patients with lower *PCAT29* expression had significantly higher rates of biochemical recurrence, consistent with our *in vitro* and *in vivo* findings.

Discussion

In this study, we characterize the novel lncRNA *PCAT29*. Our findings demonstrate that *PCAT29* is directly regulated by the androgen receptor, which binds to the promoter of *PCAT29* and suppresses its transcription. *In vitro* studies show that *PCAT29* negatively regulates prostate cancer proliferation and migration, and CAM assays demonstrate that *PCAT29* inhibits tumor growth and metastases. Low expression of *PCAT29* is associated with higher rates of biochemical recurrence, suggesting that *PCAT29* represses oncogenic phenotypes via a tumor-suppressive role.

While previous studies have nominated and characterized lncRNAs that are dysregulated in cancer (3, 14), the majority of characterized lncRNAs, to date, has been associated with oncogenic roles instead of tumor suppressor functions. In fact, there have been only a handful of lncRNAs identified to date that function in repression of cancer phenotypes, and, to our knowledge, none of these are targets downregulated by known oncogenes (14). A recent study identifies a protein-coding gene, *CCN3/NOV*, as a tumor suppressor that is repressed by androgen receptor (15). Thus, our study represents the first identification of an androgen receptor-repressed lncRNA functioning as a tumor suppressor. While further studies will be required to identify the mechanism of *PCAT29* and other tumor suppressor lncRNAs, it is clear that these lncRNAs represent an intriguing area for exploration in cancer biology.

In the context of prostate cancer, androgen-regulated lncRNAs are of fundamental importance, given that all stages of prostate cancer are exquisitely dependent on androgen receptor signaling for growth and survival. Because the majority of clinically relevant prostate cancer therapies target the androgen receptor, our studies would suggest that inhibition of androgen signaling will result in reactivation of *PCAT29*, providing another mechanism underlying the effectiveness of androgen deprivation therapy.

Clinically, there is a clear need for identification of prognostic biomarkers in prostate cancer to help guide decisions on treatment intensification. The association of high *PCAT29* expression with good clinical prognosis and

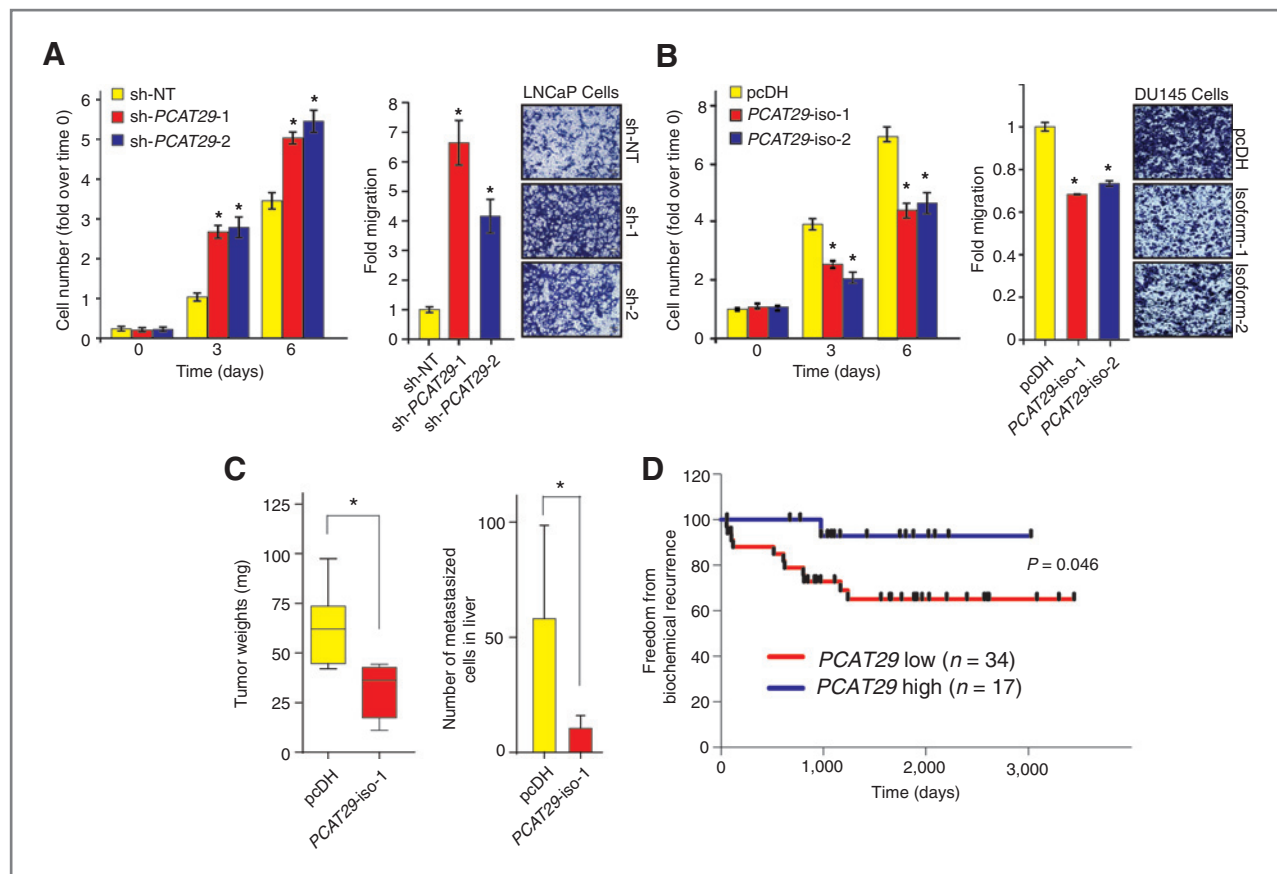


Figure 3. *PCAT29* suppresses oncogenic phenotypes. A and B, proliferation and migration of LNCaP cells stably expressing *PCAT29* shRNA (B) and DU145 cells expressing *PCAT29* expression constructs (C). Representative micrographs of crystal violet-stained migrated cells are shown as insets. C, quantification of tumor weight and metastasis to liver for 22Rv1 cells expressing *PCAT29*-isoform 1 or empty vector (pcDH) in the CAM assay. Data, mean \pm SEM. *, $P < 0.05$ by the Student *t* test. D, Kaplan–Meier analyses of prostate cancer outcomes. *PCAT29* expression was measured by qPCR and 51 patients were stratified according to their *PCAT29* expression. Patient outcomes were analyzed for freedom from biochemical recurrence.

preclinical suppression of cell proliferation and tumor metastases suggests that decreased expression or loss of *PCAT29* may identify subsets of patients who may require further intensification of therapy. As our clinical cohort was composed of hormone-sensitive disease from patients with prostatectomy, further studies need to be performed to determine whether *PCAT29* can also serve as a prognostic biomarker in the context of more advanced, castration-resistant disease. Regardless, this study highlights the importance of lncRNAs in prostate cancer biology and prognosis and suggests the need for further research in this relatively unexplored area.

Disclosure of Potential Conflicts of Interest

J.R. Prensner has ownership interest as a co-inventor on prostate cancer ncRNA patent licensed to GenomeDx Biosciences Inc., including PCATs in prostate cancer. M. Iyer has ownership interest in a patent (ncRNA and uses thereof). A.M. Chinnaiyan is a consultant/advisory board member for Wafergen Inc. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: R. Malik, D.R. Robinson, F.Y. Feng, A.M. Chinnaiyan
Development of methodology: R. Malik, L. Patel, J.R. Prensner, M.K. Iyer, I.A. Asangani, X. Jing, X. Cao, A.M. Chinnaiyan

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): R. Malik, L. Patel, J.R. Prensner, S. Subramanian, A. Carley, A. Sahu, S. Han, M. Liu, I.A. Asangani, X. Jing, X. Cao, S.M. Dhanasekaran, D.R. Robinson, F.Y. Feng

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): R. Malik, L. Patel, J.R. Prensner, Y. Shi, M.K. Iyer, A. Carley, Y.S. Niknafs, A. Sahu, I.A. Asangani, S.M. Dhanasekaran, D.R. Robinson, F.Y. Feng

Writing, review, and/or revision of the manuscript: R. Malik, J.R. Prensner, D.R. Robinson, F.Y. Feng, A.M. Chinnaiyan

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): M.K. Iyer, T. Ma, X. Jing, X. Cao, D.R. Robinson, F.Y. Feng

Study supervision: R. Malik, D.R. Robinson, F.Y. Feng, A.M. Chinnaiyan

Acknowledgments

The authors thank Xia Jiang and Vishal Kothari for helpful discussions and technical assistance, as well as Karen Giles for her review of the manuscript and submission of documents. They also thank the University of Michigan Viral vector core for generating the lentiviral constructs.

Grant Support

This work was supported in part by the NIH Prostate Specialized Program of Research Excellence Grant P50CA69568, the Early Detection Research Network grant U01 CA111275, the U.S. NIH R01CA132874-01A1, and the Department of Defense grant PC100171 (A.M. Chinnaiyan). A.M. Chinnaiyan is supported by the Doris Duke Charitable Foundation Clinical Scientist Award, the Prostate Cancer Foundation, and the Howard Hughes Medical Institute. A.M. Chinnaiyan is an American Cancer Society Research Professor and a Taubman Scholar of the

University of Michigan. F.Y. Feng was supported by the Prostate Cancer Foundation and the Department of Defense grant PC094231. R. Malik and J.R. Prensner were supported by a Prostate Cancer Foundation Young Investigator Award. R. Malik was supported by the Department of Defense Postdoctoral Fellowship W81XWH-13-1-0284. A. Sahu was supported by the NIH Predoc-

toral Fellowship 1F30CA180376-01. M. Iyer was supported by the Department of Defense Predoctoral Fellowship BC100238.

Received May 8, 2014; revised June 17, 2014; accepted June 26, 2014; published OnlineFirst July 16, 2014.

References

1. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature* 2012;489:101–8.
2. Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, et al. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat Struct Mol Biol* 2013;20:908–13.
3. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. *Cancer Discov* 2011;1:391–407.
4. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009;458:223–7.
5. Loewer S, Cabili MN, Guttman M, Loh YH, Thomas K, Park IH, et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet* 2010;42:1113–7.
6. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* 2010;329:689–93.
7. Prensner JR, Iyer MK, Sahu A, Asangani IA, Cao Q, Patel L, et al. The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nat Genet* 2013;45:1392–8.
8. Tomlins SA, Aubin SM, Siddiqui J, Lonigro RJ, Sefton-Miller L, Miick S, et al. Urine TMPRSS2:ERG fusion transcript stratifies prostate cancer risk in men with elevated serum PSA. *Sci Transl Med* 2011;3:94ra72.
9. Lee GL, Dobi A, Srivastava S. Prostate cancer: diagnostic performance of the PCA3 urine test. *Nat Rev Urol* 2011;8:123–4.
10. Asangani IA, Ateeq B, Cao Q, Dodson L, Pandhi M, Kunju LP, et al. Characterization of the EZH2-MMSET histone methyltransferase regulatory axis in cancer. *Mol Cell* 2013;49:80–93.
11. van der Horst EH, Leupold JH, Schubbert R, Ullrich A, Allgayer H. TaqMan-based quantification of invasive cells in the chick embryo metastasis assay. *Biotechniques* 2004;37:940–2, 4, 6.
12. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res* 2007;35:W345–9.
13. Yu J, Mani RS, Cao Q, Brenner CJ, Cao X, Wang X, et al. An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* 2010;17:443–54.
14. Nie L, Wu HJ, Hsu JM, Chang SS, Labaff AM, Li CW, et al. Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer. *Am J Transl Res* 2012;4:127–50.
15. Wu L, Runkle C, Jin HJ, Yu J, Li J, Yang X, et al. CCN3/NOV gene expression in human prostate cancer is directly suppressed by the androgen receptor. *Oncogene* 2014;33:504–13.

The long noncoding RNA *SChLAP1* promotes aggressive prostate cancer and antagonizes the SWI/SNF complex

John R Prensner^{1,10}, Matthew K Iyer^{1,2,10}, Anirban Sahu^{1,10}, Irfan A Asangani¹, Qi Cao¹, Lalit Patel^{1,3}, Ismael A Vergara⁴, Elai Davicioni⁴, Nicholas Erho⁴, Mercedeh Ghadessi⁴, Robert B Jenkins⁵, Timothy J Triche⁴, Rohit Malik¹, Rachel Bedenis³, Natalie McGregor³, Teng Ma⁶, Wei Chen⁶, Sumin Han⁶, Xiaojun Jing¹, Xuhong Cao¹, Xiaojun Wang¹, Benjamin Chandler¹, Wei Yan¹, Javed Siddiqui¹, Lakshmi P Kunju^{1,7,8}, Saravana M Dhanasekaran^{1,7}, Kenneth J Pienta^{1,3}, Felix Y Feng^{1,6,8} & Arul M Chinnaiyan^{1,2,7-9}

Prostate cancers remain indolent in the majority of individuals but behave aggressively in a minority^{1,2}. The molecular basis for this clinical heterogeneity remains incompletely understood³⁻⁵. Here we characterize a long noncoding RNA termed *SChLAP1* (second chromosome locus associated with prostate-1; also called *LINC00913*) that is overexpressed in a subset of prostate cancers. *SChLAP1* levels independently predict poor outcomes, including metastasis and prostate cancer-specific mortality. *In vitro* and *in vivo* gain-of-function and loss-of-function experiments indicate that *SChLAP1* is critical for cancer cell invasiveness and metastasis. Mechanistically, *SChLAP1* antagonizes the genome-wide localization and regulatory functions of the SWI/SNF chromatin-modifying complex. These results suggest that *SChLAP1* contributes to the development of lethal cancer at least in part by antagonizing the tumor-suppressive functions of the SWI/SNF complex.

With over 200,000 new cases per year, prostate cancer will be diagnosed in 1 in 6 men in the United States during their lifetime, yet only 20% of individuals with prostate cancer have a high-risk cancer that represents potentially lethal disease^{1,2,4}. Whereas mutational events in key genes characterize a subset of lethal prostate cancers^{3,5,6}, the molecular basis for aggressive disease remains poorly understood.

Long noncoding RNAs (lncRNAs) are RNA species >200 bp in length that are frequently polyadenylated and associated with transcription by RNA polymerase II (ref. 7). lncRNA-mediated biology has been implicated in a wide variety of cellular processes, and, in cancer, lncRNAs are emerging as a prominent layer of transcriptional regulation, often by collaborating with epigenetic complexes⁷⁻¹⁰.

Here we hypothesized that prostate cancer aggressiveness was governed by uncharacterized lncRNAs and sought to identify lncRNAs

associated with aggressive disease. We previously used RNA sequencing (RNA-seq) to describe 121 new lncRNA loci (out of >1,800) that were aberrantly expressed in prostate cancer tissues¹¹. Because only a fraction of prostate cancers present with aggressive clinical features², we performed cancer outlier profile analysis¹¹ (COPA) to nominate intergenic lncRNAs selectively upregulated in a subset of cancers (Supplementary Table 1). We observed that only two, *PCAT-109* and *PCAT-114*, which are both located in a 'gene desert' on chromosome 2q31.3 (Supplementary Fig. 1), had striking outlier profiles distinguishing them from the rest of the candidates¹¹ (Fig. 1a).

Of these two lncRNAs, *PCAT-114* was expressed at higher levels in prostate cell lines, and, in the *PCAT-114* region, we defined a 1.4-kb polyadenylated gene composed of up to seven exons and spanning nearly 200 kb on chromosome 2q31.3 (Fig. 1b and Supplementary Fig. 2a). We named this gene second chromosome locus associated with prostate-1 (*SChLAP1*) after its genomic location. Published prostate cancer chromatin immunoprecipitation and sequencing (ChIP-seq) data¹² confirmed that the transcriptional start site (TSS) of *SChLAP1* was marked by trimethylation of histone H3 at lysine 4 (H3K4me3) and that its gene body harbored trimethylation of histone H3 at lysine 36 (H3K36me3) (Fig. 1b), an epigenetic signature consistent with lncRNAs¹³. We observed numerous *SChLAP1* splicing isoforms, of which three (termed isoforms 1, 2 and 3) constituted the vast majority (>90%) of transcripts in the cell (Supplementary Fig. 2b,c).

Using quantitative PCR (qPCR), we confirmed that *SChLAP1* was highly expressed in ~25% of prostate cancers (Fig. 1c). *SChLAP1* was found to be expressed more frequently in metastatic compared to localized prostate cancers, and its expression was associated with *ETS* gene fusions in this cohort but not with other molecular events (Supplementary Fig. 2d,e). A computational analysis of the *SChLAP1* sequence suggested no coding potential, which was confirmed

¹Michigan Center for Translational Pathology, University of Michigan, Ann Arbor, Michigan, USA. ²Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, USA. ³Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA. ⁴GenomeDx Biosciences, Inc., Vancouver, British Columbia, Canada. ⁵Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA. ⁶Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan, USA. ⁷Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA. ⁸Comprehensive Cancer Center, University of Michigan, Ann Arbor, Michigan, USA. ⁹Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan, USA. ¹⁰These authors contributed equally to this work. Correspondence should be addressed to A.M.C. (arul@med.umich.edu).

Received 18 February; accepted 30 August; published online 29 September 2013; doi:10.1038/ng.2771

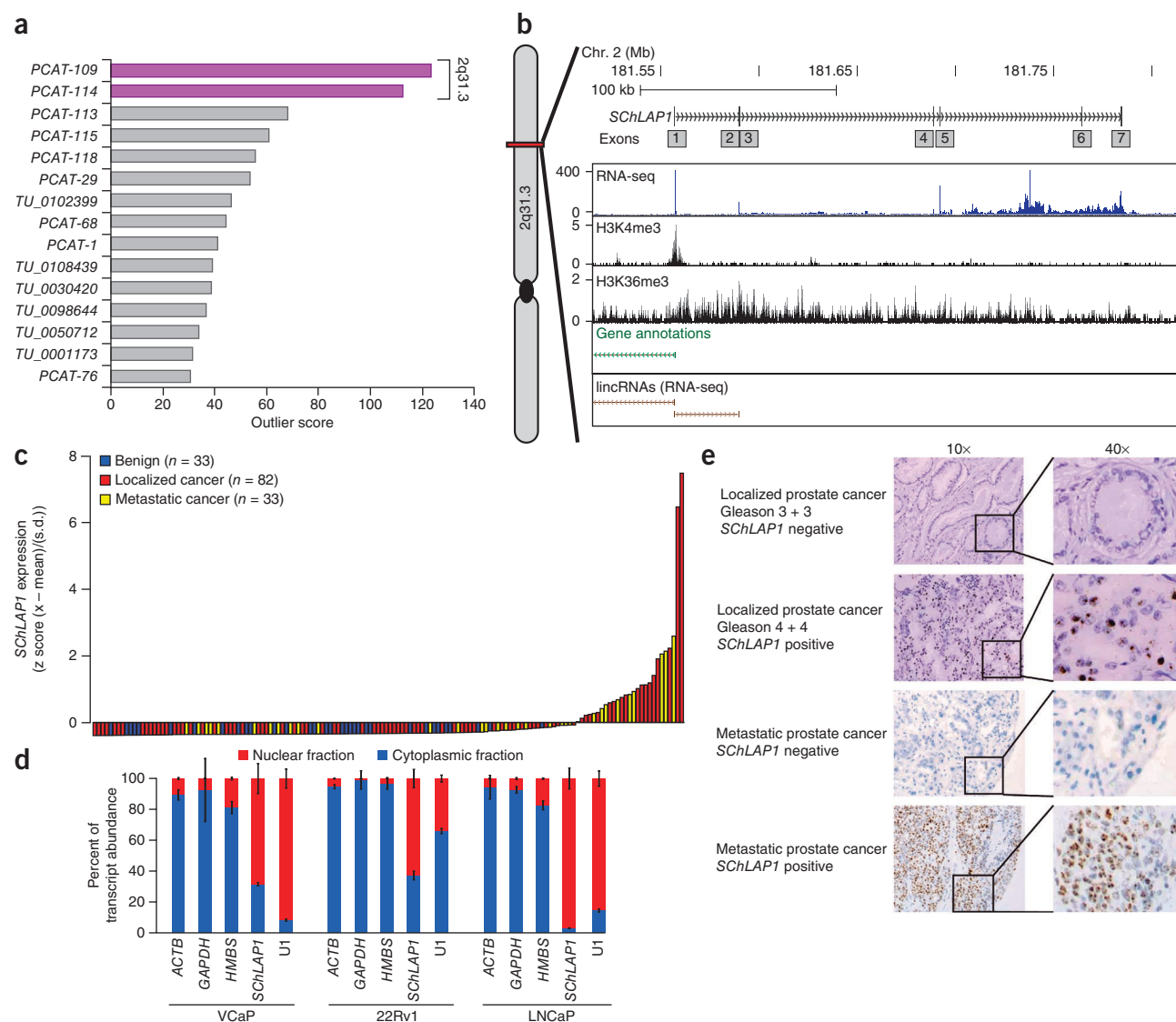


Figure 1 Identification of *SchLAP1* as a prostate cancer-associated lncRNA. (a) COPA for intergenic lncRNAs (lncRNAs defined in ref. 11). (b) Representation of the *SchLAP1* gene and its annotations in current databases. An aggregated representation of current gene annotations for Ensembl, the Encyclopedia of DNA Elements (ENCODE), the UCSC Genome Browser, RefSeq and Vega shows no annotation for *SchLAP1*. ChIP-seq data for H3K4me3 and H3K36me3 show enrichment at the *SchLAP1* gene. Also, RNA-seq data showing an outlier sample for *SchLAP1* demonstrate its expression. (c) qPCR for *SchLAP1* on a panel of benign prostate ($n = 33$), localized prostate cancer ($n = 82$) and metastatic prostate cancer ($n = 33$) samples. qPCR data are normalized to the average of (*GAPDH* + *HMBS*) and are represented as standardized expression values. (d) Fractionation of prostate cell lysates demonstrates nuclear expression of *SchLAP1*. U1 RNA serves as a positive control for nuclear gene expression. Error bars, s.e.m. (e) *In situ* hybridization of *SchLAP1* in human prostate cancer. Histological scores for localized cancer samples are indicated, with the first number representing the major Gleason score and the second number representing the minor Gleason score. *SchLAP1* staining is shown for both localized and metastatic tissues.

experimentally by *in vitro* translation assays of the three *SchLAP1* isoforms (Supplementary Fig. 3). Additionally, we found that *SchLAP1* transcripts were located in the nucleus (Fig. 1d). We confirmed the nuclear localization of *SchLAP1* transcripts in human samples (Fig. 1e) using an *in situ* hybridization assay in formalin-fixed, paraffin-embedded prostate cancer samples (Supplementary Fig. 4a,b and Supplementary Note).

An analysis of *SchLAP1* expression in localized tumors demonstrated a strong correlation with higher Gleason scores, a histopathological measure of aggressiveness (Supplementary Fig. 4c,d and Supplementary Table 2). Next, we performed a network analysis of prostate cancer microarray data in the Oncomine¹⁴ database

using signatures of *SchLAP1*-correlated or *SchLAP1*-anticorrelated genes, as *SchLAP1* itself is not measured by expression microarrays (Online Methods and Supplementary Table 3a). We found a striking association with enriched concepts related to prostate cancer progression (Fig. 2a and Supplementary Table 3b). For comparison, we next incorporated disease signatures using prostate RNA-seq data and additional known prostate cancer genes, including *EZH2* (a metastasis gene¹⁵), *PCA3* (a lncRNA biomarker⁴) and *AMACR* (a tissue biomarker⁴), as well as *ACTB* (encoding β -actin) as a control (Supplementary Fig. 5, Supplementary Table 3c–i and Supplementary Note). A heatmap visualization of significant comparisons confirmed a strong association of *SchLAP1*-correlated genes but not of

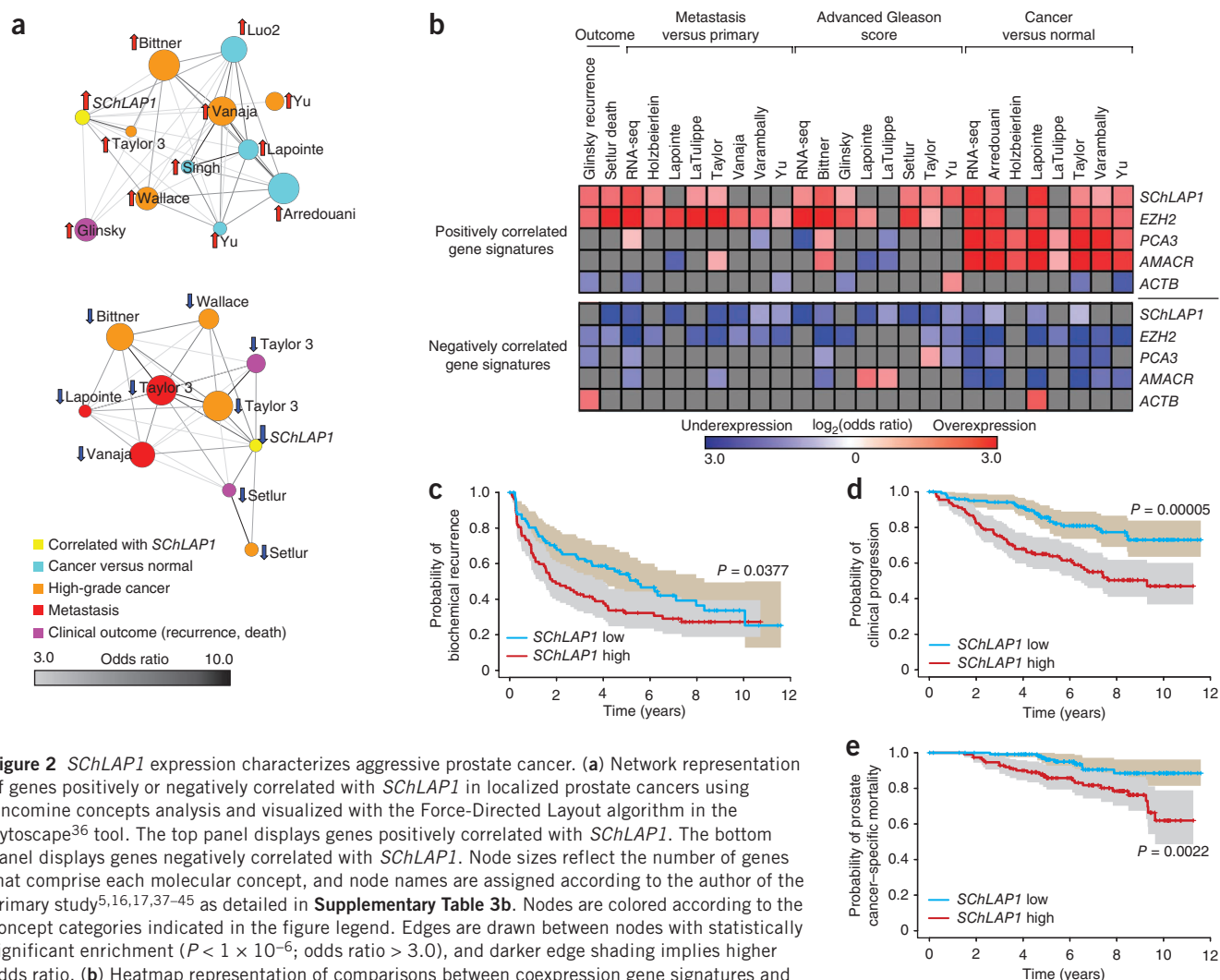


Figure 2 *SchLAP1* expression characterizes aggressive prostate cancer. **(a)** Network representation of genes positively or negatively correlated with *SchLAP1* in localized prostate cancers using OncoPrint concepts analysis and visualized with the Force-Directed Layout algorithm in the Cytoscape³⁶ tool. The top panel displays genes positively correlated with *SchLAP1*. The bottom panel displays genes negatively correlated with *SchLAP1*. Node sizes reflect the number of genes that comprise each molecular concept, and node names are assigned according to the author of the primary study^{5,16,17,37–45} as detailed in **Supplementary Table 3b**. Nodes are colored according to the concept categories indicated in the figure legend. Edges are drawn between nodes with statistically significant enrichment ($P < 1 \times 10^{-6}$; odds ratio > 3.0), and darker edge shading implies higher odds ratio. **(b)** Heatmap representation of comparisons between coexpression gene signatures and molecular concepts. Comparisons to positively (top) and negatively (bottom) correlated gene signatures are shown separately. Comparisons that do not reach statistical significance ($q > 0.01$ or odds ratio < 2) are shown in gray. Associations with overexpression concepts are colored red, and underexpression concepts are colored blue. **(c–e)** Kaplan-Meier analyses of prostate cancer outcomes in the Mayo Clinic cohort. *SchLAP1* expression was measured using Affymetrix exon arrays, and subjects were stratified according to their *SchLAP1* expression level. Subject outcomes were analyzed for biochemical recurrence (**c**), clinical progression to systemic disease (**d**) and prostate cancer-specific mortality (**e**). The shaded regions represent 95% confidence intervals. *P* values for Kaplan-Meier curves were determined using a log-rank test.

PCA3- and AMACR-correlated genes with high-grade and metastatic cancers (**Fig. 2b**). Kaplan-Meier analysis similarly showed significant associations between the *SchLAP1* signature and biochemical recurrence¹⁶ and overall survival¹⁷ (**Supplementary Fig. 6a,b**).

To directly evaluate the relationship between *SchLAP1* levels and clinical outcome, we next used *SchLAP1* expression to stratify 235 samples from individuals with localized prostate cancer who underwent radical prostatectomy at the Mayo Clinic¹⁸ (Online Methods and **Supplementary Fig. 6c**). We evaluated samples for three clinical endpoints: biochemical recurrence, clinical progression to systemic disease and prostate cancer-specific mortality (**Supplementary Table 4**). At the time of this analysis, subjects had a median follow-up time of 8.1 years.

SchLAP1 was a powerful single-gene predictor of aggressive prostate cancer (**Fig. 2c–e**). *SchLAP1* expression was highly significant when distinguishing disease with clinical progression and prostate cancer-specific mortality ($P = 0.00005$ and 0.002 , respectively; **Fig. 2d,e**). For the biochemical recurrence endpoint, high *SchLAP1*

expression was associated with a shorter median time to progression (1.9 versus 5.5 years for individuals with high and low expression of *SchLAP1*, respectively; **Fig. 2c**). We further confirmed this association with rapid biochemical recurrence using an independent cohort (**Supplementary Fig. 6d**). Multivariate and univariate regression analyses of the Mayo Clinic data demonstrated that *SchLAP1* expression is an independent predictor of prostate cancer aggressiveness, with highly significant hazard ratios for predicting biochemical recurrence, clinical progression and prostate cancer-specific mortality (hazard ratios of 3.045, 3.563 and 4.339, respectively; $P < 0.01$), which are comparable to those for other clinical factors such as advanced clinical stage and Gleason histopathological score (**Supplementary Fig. 7 and Supplementary Note**).

To explore the functional role of *SchLAP1*, we performed small interfering RNA (siRNA)-mediated knockdowns to compare the impact of *SchLAP1* depletion to that of *EZH2*, which is essential for cancer cell aggressiveness¹⁵. Notably, knockdown of *SchLAP1* dramatically impaired cell invasion and proliferation *in vitro* to

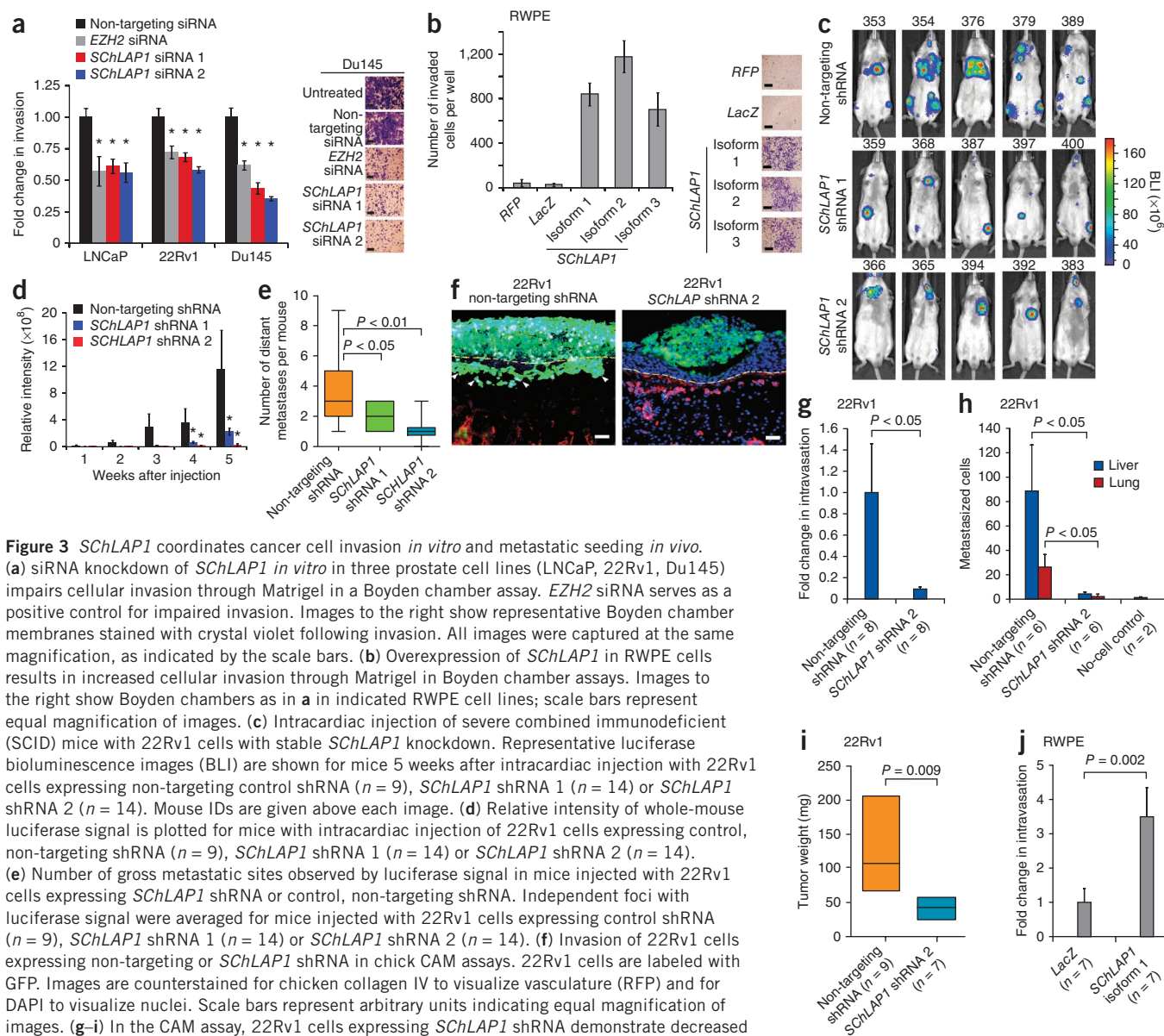


Figure 3 *SCHLAP1* coordinates cancer cell invasion *in vitro* and metastatic seeding *in vivo*.

(a) siRNA knockdown of *SCHLAP1* *in vitro* in three prostate cell lines (LNCaP, 22Rv1, Du145) impairs cellular invasion through Matrigel in a Boyden chamber assay. *EZH2* siRNA serves as a positive control for impaired invasion. Images to the right show representative Boyden chamber membranes stained with crystal violet following invasion. All images were captured at the same magnification, as indicated by the scale bars. (b) Overexpression of *SCHLAP1* in RWPE cells results in increased cellular invasion through Matrigel in Boyden chamber assays. Images to the right show Boyden chambers as in a in indicated RWPE cell lines; scale bars represent equal magnification of images. (c) Intracardiac injection of severe combined immunodeficient (SCID) mice with 22Rv1 cells with stable *SCHLAP1* knockdown. Representative luciferase bioluminescence images (BLI) are shown for mice 5 weeks after intracardiac injection with 22Rv1 cells expressing non-targeting control shRNA ($n = 9$), *SCHLAP1* shRNA 1 ($n = 14$) or *SCHLAP1* shRNA 2 ($n = 14$). Mouse IDs are given above each image. (d) Relative intensity of whole-mouse luciferase signal is plotted for mice with intracardiac injection of 22Rv1 cells expressing control, non-targeting shRNA ($n = 9$), *SCHLAP1* shRNA 1 ($n = 14$) or *SCHLAP1* shRNA 2 ($n = 14$). (e) Number of gross metastatic sites observed by luciferase signal in mice injected with 22Rv1 cells expressing *SCHLAP1* shRNA or control, non-targeting shRNA. Independent foci with luciferase signal were averaged for mice injected with 22Rv1 cells expressing control shRNA ($n = 9$), *SCHLAP1* shRNA 1 ($n = 14$) or *SCHLAP1* shRNA 2 ($n = 14$). (f) Invasion of 22Rv1 cells expressing non-targeting or *SCHLAP1* shRNA in chick CAM assays. 22Rv1 cells are labeled with GFP. Images are counterstained for chicken collagen IV to visualize vasculature (RFP) and for DAPI to visualize nuclei. Scale bars represent arbitrary units indicating equal magnification of images. (g–i) In the CAM assay, 22Rv1 cells expressing *SCHLAP1* shRNA demonstrate decreased intravasation (g), metastatic spread to the liver and lungs (h) and reduced tumor weight (i) relative to 22Rv1 cells expressing control, non-targeting shRNA. (j) Quantification of intravasation of RWPE cells expressing *LacZ* or *SCHLAP1* in the CAM assay. All data in bar plots are represented as mean \pm s.e.m. Statistical significance was determined by two-tailed Student's *t* test: $*P < 0.05$. Box plots in e,i display box-and-whisker plots with the midpoint line indicating the median, box boundaries showing 25th and 75th quartile ranges and whiskers displaying the minimum and maximum values.

an extent comparable to that observed with knockdown of *EZH2* (Fig. 3a and Supplementary Fig. 8a,b). Overexpression of an siRNA-resistant *SCHLAP1* isoform rescued the *in vitro* invasive phenotype of 22Rv1 cells treated with siRNA-2 (Supplementary Fig. 8c,d). Overexpression of the three *SCHLAP1* isoforms in benign, immortalized RWPE prostate cells dramatically increased the ability of these cells to invade *in vitro* but did not affect cell proliferation (Fig. 3b and Supplementary Fig. 8e,f).

To test *SCHLAP1* *in vivo*, we performed intracardiac injection of CB-17 SCID mice with 22Rv1 cells stably knocking down *SCHLAP1* (Supplementary Fig. 9a) and observed that *SCHLAP1* depletion impaired metastatic seeding and growth, as measured by luciferase signaling at both proximal (lungs) and distal sites (Fig. 3c,d). Indeed, compared to mice injected with 22Rv1 cells expressing a non-targeting control, mice injected with 22Rv1 cells stably expressing short

hairpin RNA (shRNA) against *SCHLAP1* had both fewer gross metastatic sites overall as well as smaller metastatic tumors when they did form (Fig. 3d,e). Histopathological analysis of the metastatic 22Rv1 tumors, regardless of *SCHLAP1* knockdown, showed uniformly high-grade epithelial cancer (Supplementary Fig. 9b). Interestingly, subcutaneous xenografts with stable knockdown of *SCHLAP1* showed slower tumor progression; however, this was due to delayed tumor engraftment rather than to decreased tumor growth kinetics, with no change in Ki67 staining observed between cells expressing *SCHLAP1* shRNA and control cells expressing non-targeting shRNA (Supplementary Fig. 9c–i).

Next, using the chick chorioallantoic membrane (CAM) assay¹⁹, we found that 22Rv1 cells expressing *SCHLAP1* shRNA 2, which have depleted expression of both isoforms 1 and 2, had greatly reduced ability to invade, intravasate and metastasize to distant organs (Fig. 3f–h). Additionally, cells with knockdown of *SCHLAP1* also

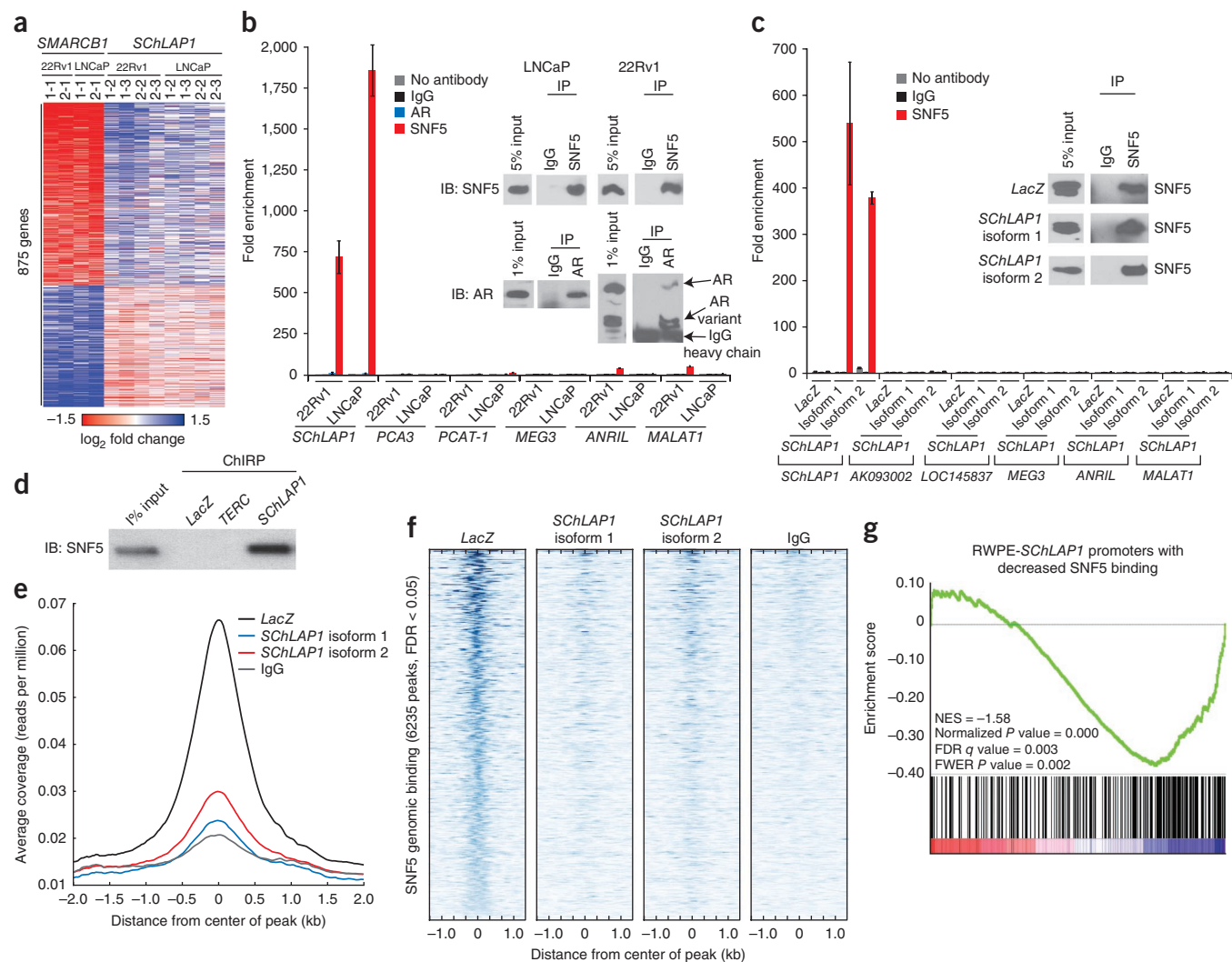


Figure 4 *SchLAP1* antagonizes SNF5 function and attenuates SNF5 genome-wide localization. (a) Heatmap results for *SchLAP1* or *SMARCB1* knockdown in LNCaP and 22Rv1 cells. The numbers above the heatmap indicate the specific siRNA and microarray replicates. (b) RIP of SNF5 or androgen receptor (AR) in 22Rv1 and LNCaP cells. Inset, protein blots showing pulldown efficiency. Error bars, s.e.m. (c) RIP analysis of SNF5 in RWPE cells overexpressing *LacZ*, *SchLAP1* isoform 1 or *SchLAP1* isoform 2. Inset, protein blots showing pulldown efficiency. Error bars, s.e.m. (d) Pulldown of *SchLAP1* RNA using chromatin isolation by RNA purification (ChIRP) recovers SNF5 protein in RWPE cells expressing *SchLAP1* isoform 1. *LacZ* and *TERC* serve as negative and positive controls, respectively. (e) Global representation of SNF5 genomic binding over a 4-kb window centered on each SNF5 ChIP-seq peak in RWPE cells expressing *LacZ*, *SchLAP1* isoform 1 or *SchLAP1* isoform 2. (f) Heatmap of SNF5 genomic binding at target sites in RWPE cells expressing *LacZ*, *SchLAP1* isoform 1 or *SchLAP1* isoform 2. A 2-kb interval centered on the called SNF5 peak is shown. (g) GSEA results showing significant enrichment of ChIP-seq promoter peaks with >2-fold loss of SNF5 binding for underexpressed genes in RWPE cells expressing *SchLAP1*. NES, normalized enrichment score; FWER, familywise error rate.

resulted in decreased tumor growth (Fig. 3i). Notably, RWPE cells with overexpression of *SchLAP1* isoform 1 partially supported these results, showing a markedly increased ability to intravasate (Fig. 3j). RWPE cells overexpressing *SchLAP1* did not generate distant metastases or cause altered tumor growth in this model (data not shown). Together, the mouse metastasis and CAM data strongly implicate *SchLAP1* in tumor invasion and metastasis through activity in cancer cell intravasation, extravasation and subsequent tumor cell seeding.

To elucidate the mechanisms of *SchLAP1* function, we profiled 22Rv1 and LNCaP cells with *SchLAP1* knockdown, identifying 165 upregulated and 264 downregulated genes (q value < 0.001) (Supplementary Fig. 10a and Supplementary Table 5a). After ranking genes according to differential expression²⁰, we employed Gene Set Enrichment Analysis (GSEA)²¹ to search for enrichment across

the Molecular Signatures Database (MSigDB)²². Among the highest ranked concepts, we noticed genes positively or negatively correlated with the SWI/SNF complex²³, and this association was independently confirmed using gene signatures generated from our RNA-seq data (Supplementary Fig. 10b–e and Supplementary Table 5b,c). Notably, *SchLAP1*-regulated genes were inversely correlated with these data sets, suggesting that *SchLAP1* and the SWI/SNF complex function in opposing manners.

The SWI/SNF complex regulates gene transcription as a multiprotein system that physically moves nucleosomes at gene promoters²⁴. Loss of SWI/SNF complex functionality promotes cancer progression, and multiple SWI/SNF components are somatically inactivated in cancer^{24,25}. SWI/SNF complex mutations do occur in prostate cancer, albeit not commonly³, and downregulation of SWI/SNF complex members

characterizes subsets of prostate cancer^{23,26}. Thus, antagonism of SWI/SNF complex activity by *SChLAP1* is consistent with the oncogenic behavior of *SChLAP1* and the tumor suppressive behavior of the SWI/SNF complex.

To directly test whether *SChLAP1* antagonizes SWI/SNF-mediated regulation, we performed siRNA-mediated knockdown of *SMARCB1* (which encodes the SNF5 protein) (Supplementary Fig. 10f), an essential subunit that facilitates SWI/SNF complex binding to histone proteins^{24,25,27}, and confirmed predicted expression changes for several *SChLAP1*- or SNF5-regulated genes (Supplementary Fig. 10g,h). A comparison of genes whose expression was altered by knockdown of *SMARCB1* to those regulated by *SChLAP1* demonstrated an antagonistic relationship in which *SChLAP1* knockdown affected the same genes as *SMARCB1* knockdown but with opposing directions of effect (Fig. 4a and Supplementary Table 5d–h). We used GSEA to quantify and verify the significance of these findings (false discovery rate (FDR) < 0.05) (Supplementary Fig. 10i–k). Furthermore, a shared *SMARCB1*-*SChLAP1* signature of coregulated genes was highly enriched for prostate cancer clinical signatures for disease aggressiveness (Supplementary Fig. 11 and Supplementary Table 5i).

Mechanistically, although *SChLAP1* and *SMARCB1* mRNA levels were comparable (Supplementary Fig. 12a), *SChLAP1* knockdown or overexpression did not alter SNF5 protein abundance (Supplementary Fig. 12b), suggesting that *SChLAP1* regulates SWI/SNF activity post-translationally. To explore this possibility, we performed RNA immunoprecipitation assays (RIPs) for SNF5. We found that endogenous *SChLAP1* but not other cytoplasmic or nuclear lncRNAs^{7,28} robustly coimmunoprecipitated with SNF5 under native conditions (Fig. 4b) and with use of UV cross-linking (Supplementary Fig. 12c), and coimmunoprecipitation was also observed with a second antibody to SNF5 (Supplementary Fig. 12d). In contrast, *SChLAP1* did not coimmunoprecipitate with androgen receptor (Fig. 4b). Furthermore, both *SChLAP1* isoform 1 and isoform 2 coimmunoprecipitated with SNF5 in RWPE overexpression models (Fig. 4c and Supplementary Fig. 12e). SNRNP70 binding to U1 RNA was used as a technical control in all cell lines (Supplementary Fig. 12f,g). Finally, pulldown of *SChLAP1* RNA in RWPE cells overexpressing *SChLAP1* isoform 1 robustly recovered SNF5 protein, confirming this interaction (Fig. 4d and Supplementary Fig. 12h).

To address whether *SChLAP1* modulates SWI/SNF genomic binding, we performed ChIP-seq for SNF5 in RWPE cells expressing *LacZ* or *SChLAP1* and called significantly enriched peaks with respect to an IgG control (Online Methods and Supplementary Table 6a). Protein blot validation confirmed SNF5 pulldown by ChIP (Supplementary Fig. 13a). After aggregating called peaks from all samples, we found 6,235 genome-wide binding sites for SNF5 (FDR < 0.05; Supplementary Table 6b), which were highly enriched for sites near gene promoters (Supplementary Fig. 13b), supporting results from previous studies of SWI/SNF binding^{29–31}.

A comparison of SNF5 binding across these 6,235 genomic sites demonstrated a dramatic decrease in SNF5 genomic binding as a result of *SChLAP1* overexpression (Fig. 4e,f and Supplementary Fig. 13c). Of the 1,299 SNF5 peaks occurring within 1 kb of a gene TSS, 390 showed relative SNF5 binding that was decreased by ≥ 2 -fold with *SChLAP1* overexpression (Supplementary Fig. 13d and Supplementary Table 6c). To verify these findings independently, we performed ChIP for SNF5 in 22Rv1 cells expressing shRNA to *SChLAP1*, with the hypothesis that knockdown of *SChLAP1* should increase SNF5 genomic binding compared to controls. We found that 9 of 12 target genes showed a substantial increase in SNF5 binding with knockdown of *SChLAP1* (Supplementary Fig. 14a), confirming our predictions.

Finally, we used expression profiling of RWPE cells expressing *LacZ* or *SChLAP1* to characterize the relationship between SNF5 binding and *SChLAP1*-mediated changes in gene expression. After identifying a gene signature with highly significant changes in expression (Supplementary Table 6d), we intersected this signature with the ChIP-seq data. We observed that a substantial subset of genes with ≥ 2 -fold relative decrease in SNF5 genomic binding were dysregulated when *SChLAP1* was overexpressed (Supplementary Fig. 14b). Decreased SNF5 binding was primarily associated with the down-regulation of target gene expression (Supplementary Table 6e), although the SWI/SNF complex is known to regulate expression in either direction^{24,25}. Integrative GSEA of the microarray and SNF5 ChIP-seq data demonstrated significant enrichment for genes that were repressed when *SChLAP1* was overexpressed (q value = 0.003; Fig. 4g). Overall, these data argue that *SChLAP1* overexpression antagonizes SWI/SNF complex function by attenuating the genomic binding of this complex, thereby impairing its ability to properly regulate gene expression.

Here we have discovered *SChLAP1*, a highly prognostic lncRNA that is abundantly expressed in $\sim 25\%$ of prostate cancers and that aids in the discrimination of aggressive tumors from indolent forms of the disease. Mechanistically, we find that *SChLAP1* coordinates cancer cell invasion *in vitro* and metastatic spread *in vivo*. Moreover, we characterize an antagonistic *SChLAP1*-SWI/SNF axis in which *SChLAP1* impairs SNF5-mediated regulation of gene expression and genomic binding (Supplementary Fig. 14c). Thus, whereas other lncRNAs such as *HOTAIR* and *HOTTIP* are known to assist epigenetic complexes such as PRC2 and MLL by facilitating their genomic binding and enhancing their functions^{8,9,32}, *SChLAP1* is the first lncRNA, to our knowledge, that impairs a major epigenetic complex with well-documented tumor suppressor function^{23–25,33–35}. Our discovery of *SChLAP1* has broad implications for cancer biology and provides supporting evidence for the role of lncRNAs in the progression of aggressive cancers.

URLs. Stellaris probe designer, <http://www.singlemoleculefish.com>; HT-Seq, <http://www-huber.embl.de/users/anders/HTSeq/>; BioVenn, <http://www.cmbi.ru.nl/cdd/biovenn/>; Galaxy, <http://usegalaxy.org/>.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Sequences for *SChLAP1* isoforms 1–7 have been deposited in GenBank under accessions [JX117418](#), [JX117419](#), [JX117420](#), [JX117421](#), [JX117422](#), [JX117423](#) and [JX117424](#). Microarray data have been deposited in the Gene Expression Omnibus (GEO) under accession [GSE40386](#).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank O.A. Balbin, S.A. Tomlins, C. Brenner, S. Deroo and S. Roychowdhury for helpful discussions. This work was supported in part by US National Institutes of Health (NIH) Prostate Specialized Program of Research Excellence grant P50CA69568, Early Detection Research Network grant UO1 CA111275, US NIH grant R01CA132874-01A1 and US Department of Defense grant PC100171 (A.M.C.). A.M.C. is supported by a Doris Duke Charitable Foundation Clinical Scientist Award, by the Prostate Cancer Foundation and by the Howard Hughes Medical Institute. A.M.C. is an American Cancer Society Research Professor. A.M.C. is a Taubman Scholar of the University of Michigan. F.Y.F. was supported by the Prostate Cancer Foundation and by US Department of Defense grant PC094231. Q.C. was supported by US Department of Defense Postdoctoral

Fellowship PC094725. J.R.P. was supported by US Department of Defense Predoctoral Fellowship PC094290. M.K.I. was supported by US Department of Defense Predoctoral Fellowship BC100238. A.S. was supported by NIH Predoctoral Fellowship 1F30CA180376-01. J.R.P., M.K.I. and A.S. are Fellows of the University of Michigan Medical Scientist Training Program.

AUTHOR CONTRIBUTIONS

J.R.P., M.K.I., A.S. and A.M.C. designed the project and directed experimental studies. J.R.P., Q.C., W.C., S.M.D., B.C., S.H., R.M., L.P., T.M. and A.S. performed *in vitro* studies. X.W. performed *in vitro* translation assays. I.A.A. and A.S. performed CAM assays. R.B., N.M. and K.J.P. performed *in vivo* studies. L.P.K. and W.Y. performed histopathological analyses. M.K.I. performed bioinformatics analysis. X.J. and X.C. performed gene expression microarray experiments. J.S. and F.Y.F. facilitated biological sample procurement. F.Y.F. performed clinical analyses. For the Mayo Clinic cohort, R.B.J. provided clinical samples and outcomes data. T.J.T. and E.D. generated and analyzed expression profiles for the Mayo Clinic cohort. E.D., N.E., M.G. and I.A.V. performed statistical analyses of *SchLAPI* expression in the Mayo Clinic cohort. J.R.P., M.K.I., A.S. and A.M.C. interpreted data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the [online version of the paper](#).

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Etzioni, R., Cha, R., Feuer, E.J. & Davidov, O. Asymptomatic incidence and duration of prostate cancer. *Am. J. Epidemiol.* **148**, 775–785 (1998).
- Cooperberg, M.R., Moul, J.W. & Carroll, P.R. The changing face of prostate cancer. *J. Clin. Oncol.* **23**, 8146–8151 (2005).
- Grasso, C.S. *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012).
- Prensner, J.R., Rubin, M.A., Wei, J.T. & Chinnaiyan, A.M. Beyond PSA: the next generation of prostate cancer biomarkers. *Sci. Transl. Med.* **4**, 127rv3 (2012).
- Taylor, B.S. *et al.* Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**, 11–22 (2010).
- Berger, M.F. *et al.* The genomic complexity of primary human prostate cancer. *Nature* **470**, 214–220 (2011).
- Prensner, J.R. & Chinnaiyan, A.M. The emergence of lncRNAs in cancer biology. *Cancer Discov.* **1**, 391–407 (2011).
- Rinn, J.L. *et al.* Functional demarcation of active and silent chromatin domains in human *HOX* loci by noncoding RNAs. *Cell* **129**, 1311–1323 (2007).
- Tsai, M.C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689–693 (2010).
- Kotake, Y. *et al.* Long non-coding RNA *ANRIL* is required for the PRC2 recruitment to and silencing of *p15^{INK4B}* tumor suppressor gene. *Oncogene* **30**, 1956–1962 (2011).
- Prensner, J.R. *et al.* Transcriptome sequencing across a prostate cancer cohort identifies *PCAT-1*, an unannotated lincRNA implicated in disease progression. *Nat. Biotechnol.* **29**, 742–749 (2011).
- Yu, J. *et al.* An integrated network of androgen receptor, polycomb, and *TMPRSS2-ERG* gene fusions in prostate cancer progression. *Cancer Cell* **17**, 443–454 (2010).
- Guttman, M. *et al.* Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**, 223–227 (2009).
- Rhodes, D.R. *et al.* Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* **9**, 166–180 (2007).
- Varambally, S. *et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature* **419**, 624–629 (2002).
- Glinsky, G.V., Glinskii, A.B., Stephenson, A.J., Hoffman, R.M. & Gerald, W.L. Gene expression profiling predicts clinical outcome of prostate cancer. *J. Clin. Invest.* **113**, 913–923 (2004).
- Settlur, S.R. *et al.* Estrogen-dependent signaling in a molecularly distinct subclass of aggressive prostate cancer. *J. Natl. Cancer Inst.* **100**, 815–825 (2008).
- Nakagawa, T. *et al.* A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy. *PLoS ONE* **3**, e2318 (2008).
- Asangani, I.A. *et al.* Characterization of the EZH2-MMSET histone methyltransferase regulatory axis in cancer. *Mol. Cell* **49**, 80–93 (2013).
- Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121 (2001).
- Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
- Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
- Shen, H. *et al.* The SWI/SNF ATPase Brm is a gatekeeper of proliferative control in prostate cancer. *Cancer Res.* **68**, 10154–10162 (2008).
- Roberts, C.W. & Orkin, S.H. The SWI/SNF complex—chromatin and cancer. *Nat. Rev. Cancer* **4**, 133–142 (2004).
- Reisman, D., Glaros, S. & Thompson, E.A. The SWI/SNF complex and cancer. *Oncogene* **28**, 1653–1668 (2009).
- Sun, A. *et al.* Aberrant expression of SWI/SNF catalytic subunits BRG1/BRM is associated with tumor development and increased invasiveness in prostate cancers. *Prostate* **67**, 203–213 (2007).
- Dechassa, M.L. *et al.* Architecture of the SWI/SNF-nucleosome complex. *Mol. Cell. Biol.* **28**, 6010–6021 (2008).
- Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789 (2012).
- De, S. *et al.* Dynamic BRG1 recruitment during T helper differentiation and activation reveals distal regulatory elements. *Mol. Cell. Biol.* **31**, 1512–1527 (2011).
- Euskirchen, G.M. *et al.* Diverse roles and interactions of the SWI/SNF chromatin remodeling complex revealed using global approaches. *PLoS Genet.* **7**, e1002008 (2011).
- Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T. & Pugh, B.F. Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* **149**, 1461–1473 (2012).
- Gupta, R.A. *et al.* Long non-coding RNA *HOTAIR* reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
- Jones, S. *et al.* Frequent mutations of chromatin remodeling gene *ARID1A* in ovarian clear cell carcinoma. *Science* **330**, 228–231 (2010).
- Varela, I. *et al.* Exome sequencing identifies frequent mutation of the SWI/SNF complex gene *PBRM1* in renal carcinoma. *Nature* **469**, 539–542 (2011).
- Versteeg, I. *et al.* Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**, 203–206 (1998).
- Cline, M.S. *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc.* **2**, 2366–2382 (2007).
- Arredouani, M.S. *et al.* Identification of the transcription factor single-minded homologue 2 as a potential biomarker and immunotherapy target in prostate cancer. *Clin. Cancer Res.* **15**, 5794–5802 (2009).
- Holzbeierlein, J. *et al.* Gene expression analysis of human prostate carcinoma during hormonal therapy identifies androgen-responsive genes and mechanisms of therapy resistance. *Am. J. Pathol.* **164**, 217–227 (2004).
- Lapointe, J. *et al.* Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc. Natl. Acad. Sci. USA* **101**, 811–816 (2004).
- LaTulippe, E. *et al.* Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease. *Cancer Res.* **62**, 4499–4506 (2002).
- Luo, J.H. *et al.* Gene expression analysis of prostate cancers. *Mol. Carcinog.* **33**, 25–35 (2002).
- Vanaja, D.K., Chevillat, J.C., Iturria, S.J. & Young, C.Y. Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression. *Cancer Res.* **63**, 3877–3882 (2003).
- Varambally, S. *et al.* Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* **8**, 393–406 (2005).
- Wallace, T.A. *et al.* Tumor immunobiological differences in prostate cancer between African-American and European-American men. *Cancer Res.* **68**, 927–936 (2008).
- Yu, Y.P. *et al.* Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. *J. Clin. Oncol.* **22**, 2790–2799 (2004).

ONLINE METHODS

Cell lines. All cell lines were obtained from the American Type Culture Collection. Cell lines were maintained using standard media and conditions. Specifically, VCaP and Du145 cells were maintained in DMEM (Invitrogen) supplemented with 10% FBS and 1% penicillin-streptomycin. LNCaP and 22Rv1 cells were maintained in RPMI 1640 (Invitrogen) supplemented with 10% FBS and 1% penicillin-streptomycin. RWPE cells were maintained in KSF medium (Invitrogen) supplemented with 10 ng/ml epidermal growth factor (EGF; Sigma) and bovine pituitary extract (BPE) and with 1% penicillin-streptomycin. All cell lines were grown at 37 °C in a 5% CO₂ cell culture incubator. All cell lines were genotyped for identity at the University of Michigan Sequencing Core and were tested routinely for Mycoplasma contamination.

Cell lines expressing *SchLAP1* or control constructs were generated by cloning *SchLAP1* or control sequence into the pLenti6 vector (Invitrogen), using pcr8 non-directional Gateway cloning (Invitrogen) as an initial cloning vector, and shuttling to pLenti6 using LR clonase II (Invitrogen) according to the manufacturer's instructions. Stably transfected RWPE and 22Rv1 cells were selected with blasticidin (Invitrogen) for 1 week. For LNCaP and 22Rv1 cells with stable knockdown of *SchLAP1*, cells were transfected with lentiviral constructs encoding *SchLAP1* shRNA or with non-targeting shRNA lentiviral constructs for 48 h. GFP-positive cells were selected with 1 µg/ml puromycin for 72 h. All lentiviruses were generated by the University of Michigan Vector Core.

Tissue samples. Prostate tissues were obtained from the radical prostatectomy series and Rapid Autopsy Program at the University of Michigan tissue core⁴⁶. These programs are part of the University of Michigan Prostate Cancer Specialized Program of Research Excellence (SPORE). All tissue samples were collected with informed consent under an institutional review board (IRB)-approved protocol at the University of Michigan (SPORE in Prostate Cancer (Tissue/Serum/Urine) Bank Institutional Review Board 1994-0481).

RNA isolation and cDNA synthesis. Total RNA was isolated using TRIzol (Invitrogen) and an RNeasy kit (Qiagen) with DNase I digestion according to the manufacturers' instructions. RNA integrity was verified on an Agilent Bioanalyzer 2100 (Agilent Technologies). cDNA was synthesized from total RNA using Superscript III (Invitrogen) and random primers (Invitrogen).

Quantitative RT-PCR. Quantitative RT-PCR was performed using Power SYBR Green MasterMix (Applied Biosystems) on an Applied Biosystems 7900HT Real-Time PCR System. All oligonucleotide primers were obtained from Integrated DNA Technologies (IDT), and primer sequences are listed in **Supplementary Table 7a**. The housekeeping genes *GAPDH*, *HMBS* and *ACTB* were used as loading controls. Fold changes were calculated relative to housekeeping genes and were normalized to the median value in benign samples.

RT-PCR. RT-PCR was performed for primer pairs using Platinum Taq High-Fidelity polymerase (Invitrogen). PCR products were resolved on a 1.0% agarose gel. PCR products were then either sequenced directly (if only a single product was observed) or appropriate gel products were extracted using a Gel Extraction kit (Qiagen) and cloned into pcr4-TOPO vector (Invitrogen). PCR products were bidirectionally sequenced at the University of Michigan Sequencing Core using either gene-specific primers or M13 forward and reverse primers for cloned PCR products. All oligonucleotide primers were obtained from IDT, and primer sequences are listed in **Supplementary Table 7a**.

RACE. 5' and 3' RACE were performed using the GeneRacer RLM-RACE kit (Invitrogen) according to the manufacturer's instructions. RACE PCR products were obtained using Platinum Taq High-Fidelity polymerase, the supplied GeneRacer primers and the appropriate gene-specific primers indicated in **Supplementary Table 7a**. RACE PCR products were separated on a 1.5% agarose gel. Gel products were extracted with a Gel Extraction kit, cloned into pcr4-TOPO vectors and sequenced bidirectionally using M13 forward and reverse primers at the University of Michigan Sequencing Core. At least three colonies were sequenced for every RACE PCR product that was gel purified.

siRNA-mediated knockdown. Cells were plated in 100-mm plates at a desired concentration and transfected with 20 µM experimental siRNA oligonucleotides or non-targeting controls twice at 8 h and 24 h after plating. Knockdown was performed with Oligofectamine in OptiMEM medium. Knockdown efficiency was determined by qPCR. siRNA sequences (in sense orientation) for knockdown experiments are listed in **Supplementary Table 7b**. At 72 h after transfection, cells were trypsinized, counted with a Coulter counter and diluted to 1 million cells/ml.

Overexpression. Full-length *SchLAP1* transcript was amplified from LNCaP cells and cloned into the pLenti6 vector along with *LacZ* control sequence. Insert sequences were confirmed by Sanger sequencing at the University of Michigan Sequencing Core. Lentiviruses were generated at the University of Michigan Vector Core. The benign immortalized prostate cell line RWPE was infected with lentiviruses expressing *SchLAP1* or *LacZ*, and stable pools and clones were generated by selection with blasticidin. Similarly, the immortalized cancer cell line 22Rv1 was infected with lentiviruses expressing *SchLAP1* or *LacZ*, and stable pools were generated by selection with blasticidin.

Cell proliferation assays. At 72 h after transfection with siRNA, cells were trypsinized, counted with a Coulter counter and diluted to 1 million cells/ml. For proliferation assays, 10,000 cells were plated in each well of a 24-well plate and grown in regular growth medium. At 48 h and 96 h after plating, cells were collected by trypsinizing and counted using a Coulter counter. All assays were performed in quadruplicate.

Basement membrane matrix invasion assays. For invasion assays, cells were treated with the indicated siRNAs, and, at 72 h after transfection, cells were trypsinized, counted with a Coulter counter and diluted to 1 million cells/ml. Cells were seeded onto basement membrane matrix (EC matrix, Chemicon) present in the insert of a 24-well culture plate. FBS was added to the lower chamber as a chemoattractant. After 48 h, the non-invading cells and EC matrix were gently removed with a cotton swab. Invasive cells located on the lower side of the chamber were stained with crystal violet, air dried and photographed. For colorimetric assays, inserts were treated with 150 µl of 10% acetic acid, and absorbance was measured at 560 nm using a spectrophotometer (GE Healthcare).

shRNA-mediated knockdown. The prostate cancer cell lines LNCaP and 22Rv1 were seeded at 50–60% confluency and were allowed to attach overnight. Cells were transfected with lentiviral constructs expressing *SchLAP1* or non-targeting shRNA as described previously for 48 h. GFP-positive cells were selected with 1 µg/ml puromycin for 72 h. At 48 h after the start of selection, cells were collected for protein and RNA using RIPA buffer or TRIzol, respectively. RNA was processed as described above.

Gene expression profiling. Expression profiling was performed using the Agilent Whole Human Genome Oligo Microarray according to previously published protocols⁴⁷. All samples were run in technical triplicates, comparing knockdown samples treated with *SchLAP1* siRNA to samples treated with non-targeting control siRNA. Expression data were analyzed using the SAM method as described previously²⁰.

Mouse intracardiac and subcutaneous *in vivo* models. All experimental procedures were approved by the University of Michigan Committee for the Use and Care of Animals (UCUCA).

For the intracardiac injection model, 5×10^5 cells from 1 of 3 experimental cell lines (22Rv1-sh*SchLAP1*-1 or 22Rv1-sh*SchLAP1*-2 (two cell lines expressing *SchLAP1* shRNA) or 22Rv1-shNT (expressing control vector), all with luciferase constructs incorporated) were introduced into CB-17 SCID mice at 6 weeks of age. Female mice were used to minimize endogenous androgen production that might stimulate xenografted prostate cells. We used 15 mice per cell line to ensure adequate statistical power to distinguish phenotypes between groups. Mice used in these studies were randomized by double-blind injection of cell line samples into mice and were monitored for tumor growth by researchers blinded to the study design. Beginning 1 week after injection, bioluminescent imaging of mice was performed weekly using a CCD IVIS

system with a 50-mm lens (Xenogen), and the results were analyzed using LivingImage software (Xenogen). When the mouse reached the determined end point, defined as whole-body region of interest (ROI) of 1×10^{10} photons, or became fatally ill, it was euthanized, and the lung and liver were resected. Half of the resected specimen was placed in an immunohistochemistry cassette, incubated in 10% buffered formalin phosphate (Fisher Scientific) for 24 h and transferred to 70% ethanol until further analysis. The other half of each specimen was snap frozen in liquid nitrogen and stored at -80°C . A specimen was disregarded if the tumor was localized only in the heart. After accounting for these considerations, there were 9 mice analyzed for 22Rv1-shNT cells and 14 mice each analyzed for 22Rv1-sh*SchLAP1*-1 and 22Rv1-sh*SchLAP1*-2 cells.

For the subcutaneous injection model, 1×10^6 cells from 1 of the 3 previously described experimental cell lines were introduced into mice (CB-17 SCID), aged 5–7 weeks, with a Matrigel scaffold (BD Matrigel Matrix, BD Biosciences) in the posterior dorsal flank region ($n = 10$ per cell line). Tumors were measured weekly using a digital caliper, and the end point was defined by tumor volume of $1,000\text{ mm}^3$. When a mouse reached the end point or became fatally ill, it was euthanized, and the primary tumor was resected. The resected specimen was divided in half: one half was placed in 10% buffer formalin, and the other half was snap frozen. For histological analyses, formalin-fixed, paraffin-embedded mouse livers and lungs were sectioned on a microtome into $5\text{-}\mu\text{m}$ sections on glass slides. Slides were stained with hematoxylin and eosin using standard methods and were analyzed by a board-certified pathologist (L.P.K.).

Immunoblot analysis. Cells were lysed in RIPA lysis buffer (Sigma) supplemented with HALT protease inhibitor (Fisher). Protein blotting analysis was performed with standard protocols using polyvinylidene difluoride (PVDF) membrane (GE Healthcare), and signals were visualized with an enhanced chemiluminescence system as described by the manufacturer (GE Healthcare).

Protein lysates were boiled in sample buffer, and $10\text{ }\mu\text{g}$ of protein was loaded onto an SDS-PAGE gel and run for separation of proteins. Proteins were transferred onto PVDF membrane and blocked for 90 min in blocking buffer (5% milk in a solution of 0.1% Tween-20 in Tris-buffered saline (TBS-T)). Membranes were incubated overnight at 4°C with primary antibody. After three washes with TBS-T and one wash with TBS, the blot was incubated with HRP-conjugated secondary antibody, and signal was visualized with an enhanced chemiluminescence system as described by the manufacturer. Primary antibodies used included antibody to SNF5 (1:1,000 dilution; Millipore, ABD22, rabbit), SNF5 (1:1,000 dilution; Abcam, ab58209, mouse), β -actin (1:5,000 dilution; Sigma, A5316, mouse) and androgen receptor (1:1,000 dilution; Millipore, 06-680, rabbit).

RIP assays. RIP assays were performed using a Millipore EZ-Magna RIP RNA-Binding Protein Immunoprecipitation kit (Millipore, 17-701) according to the manufacturer's instructions. RIP PCR was performed as qPCR, as described above, using total RNA as input controls. We used 1/150 volume of the RIP RNA product per PCR reaction. Antibodies used for RIP included rabbit polyclonal IgG (Millipore, PP64) and antibodies to SNRNP70 (Millipore, CS203216), SNF5 (Millipore, ABD22, rabbit), SNF5 (Abcam, ab58209, mouse) and androgen receptor (Millipore, 06-680, rabbit), and 5–7 μg of antibody was used per RIP reaction. All RIP assays were performed in biological duplicate. For UV-crosslinked RIP experiments, cells were subjected to 400 J of 254 nm UV light twice and were then collected for RIP experiments as described above.

ChIP assays. ChIP assays were performed as described previously^{11,12} using antibody for SNF5 (Millipore, ABD22, rabbit) and rabbit IgG (Millipore, PP64B). Briefly, approximately 1 million cells were cross-linked per antibody for 10–15 min with 1% formaldehyde, and crosslinking was inactivated by incubation with 0.125 M glycine for 5 min at room temperature. Cells were rinsed with cold PBS three times, and cell pellets were resuspended in lysis buffer supplemented with protease inhibitors. Chromatin was sonicated to an average length of 500 bp and centrifuged to remove debris, and supernatants containing chromatin fragments were incubated with protein A or protein G

beads to reduce non-specific binding. Beads were then removed, and supernatants were incubated with 6 μg of antibody overnight at 4°C . Fresh beads were added and incubated with protein-chromatin-antibody complexes for 2 h at 4°C , washed twice with $1\times$ dialysis buffer and four times with IP wash buffer, and eluted in 150 μl of IP elution buffer¹². One-tenth of the ChIP reaction was taken for protein evaluation for validation of pulldown. Cross-linking was reversed by incubating eluted products with 0.3 M NaCl at 65°C overnight. ChIP products were cleaned with the USB PrepEase kit. ChIP experiments were validated for specificity of the antibody by protein blotting.

ChIP-seq experiments. Paired-end ChIP-seq libraries were generated following the Illumina ChIP-seq protocol with minor modifications. DNA isolated by ChIP assay was subjected to end repair and A tailing before ligation with Illumina adaptors. Samples were purified using AMPure beads (Beckman Coulter) and PCR enriched with a combination of specific index primers and PE2.0 primer under the following conditions: 98°C (30 s), 65°C (30 s) and 72°C (40 s, with the addition of 4 s per cycle). After 14 cycles of amplification, a final extension at 72°C for 5 min was carried out. Barcoded libraries were size selected using 3% NuSieve Agarose gels (Lonza) and subjected to an additional PCR enrichment step. Libraries were analyzed and quantified using a Bioanalyzer instrument (Agilent Technologies) before they were subjected to paired-end sequencing using the Illumina HiSeq platform.

CAM assays. CAM assays were performed as previously described¹⁹. Briefly, fertilized chicken eggs were incubated in a rotary humidified incubator at 38°C for 10 d. CAM was released by applying a mild amount of pressure to the hole over the air sac and cutting a 1-cm^2 window encompassing a second hole near the allantoic vein. Approximately 2 million cells in 50 μl of medium were implanted in each egg, windows were sealed, and eggs were returned to a stationary incubator.

For local invasion and intravasation experiments, the upper and lower CAMs were isolated after 72 h. Upper CAMs were processed and stained for chicken collagen IV (immunofluorescence) or human cytokeratin (immunohistochemistry) as previously described¹⁹.

For metastasis assays, embryonic livers were isolated on day 18 of embryonic growth and analyzed for the presence of tumor cells by quantitative human Alu-specific PCR. Genomic DNA isolates from lower CAMs and livers were prepared using the Puregene DNA purification system (Qiagen), and quantification with human Alu-specific PCR was performed as described¹⁹. Fluorogenic TaqMan qPCR probes were generated as described above and used to determine DNA copy number.

For xenograft growth assays with RWPE cells, embryos were sacrificed on day 18, and extraembryonic xenografts were excised and weighed.

In situ hybridization. *In situ* hybridization assays were performed as a commercial service from Advanced Cell Diagnostics, Inc. Briefly, cells in the clinical specimens were fixed and permeabilized using xylene, ethanol and protease to allow for probe access. Slides were boiled in pretreatment buffer for 15 min and rinsed in water. Next, two independent target probes were hybridized to *SchLAP1* RNA at 40°C for 2 h, with this pair of probes creating a binding site for a preamplifier. After this incubation, the preamplifier was hybridized to the target probes at 30°C and amplified with six cycles of hybridization followed by two washes. Cells were counterstained to visualize signal. Finally, slides were stained with hematoxylin and eosin, dehydrated with 100% ethanol and xylene and mounted in a xylene-based mounting medium.

In vitro translation. Full-length *SchLAP1*, *PCAT-1* or *GUS* positive control sequences were cloned into the PCR2.1 entry vector (Invitrogen). Insert sequences were confirmed by Sanger sequencing at the University of Michigan Sequencing Core. *In vitro* translation assays were performed with the TnT Quick Coupled Transcription/Translation System (Promega) with 1 mM methionine and Transcend Biotin-Lysyl-tRNA (Promega) according to the manufacturer's instructions.

ChIRP assays. ChIRP assays were performed as previously described⁴⁸. Briefly, antisense DNA probes targeting the full-length *SchLAP1* sequence

were designed using the online designer at Stellaris (see URLs). Fifteen probes spanning the entire transcript and unique to the *SchLAP1* sequence were chosen. Additionally, ten probes were designed against *TERC* RNA as a positive control, and 24 probes were designed against *LacZ* RNA as a negative control. All probes were synthesized with 3' biotinylation (IDT). Sequences of all probes are listed in **Supplementary Table 8**. RWPE cells overexpressing *SchLAP1* isoform 1 were grown to 80% confluency in 100-mm cell culture dishes. Two dishes were used for each probe set. Before being collected, cells were rinsed with 1× PBS and cross-linked with 1% glutaraldehyde (Sigma) for 10 min at room temperature. Cross-linking was quenched by incubation with 0.125 M glycine for 5 min at room temperature. Cells were rinsed twice with 1× PBS, collected and pelleted at 1,500g for 5 min. Nuclei were isolated using the Pierce NE-PER Nuclear Protein Extraction kit. Nuclear pellets were resuspended in 100 mg/ml cell lysis buffer (50 mM Tris, pH 7.0, 10 mM EDTA, 1% SDS, and, added before use, 1 mM dithiothreitol (DTT), phenylmethylsulfonyl fluoride (PMSF), protease inhibitor and Superscript-III (Invitrogen)). Lysates were placed on ice for 10 min and sonicated using a Bioruptor (Diagenode) at the highest setting with 30-s on and 45-s off cycles until lysates were completely solubilized. Cell lysates were diluted in twice the volume of hybridization buffer (500 mM NaCl, 1% SDS, 100 mM Tris, pH 7.0, 10 mM EDTA, 15% formamide, and, added before use, DTT, PMSF, protease inhibitor and Superscript-III), and 100 nM probes were added to the diluted lysates. Hybridization was carried out by end-over-end rotation at 37 °C for 4 h. Magnetic streptavidin C1 beads were prepared by washing three times in cell lysis buffer and were then added to each hybridization reaction at a concentration of 100 µl per 100 pmol of probe. Reactions were incubated at 37 °C for 30 min with end-over-end rotation. Bead-probe-RNA complexes were captured with magnetic racks (Millipore) and washed five times with 1 ml wash buffer (2× SSC, 0.5% SDS, fresh PMSF added). After the final wash, 20% of the sample was used for RNA isolation, and 80% of the sample was used for protein isolation. For RNA elution, beads were resuspended in 200 µl of RNA proteinase K buffer (100 mM NaCl, 10 mM Tris, pH 7.0, 1 mM EDTA, 0.5% SDS) and 1 mg/ml proteinase K (Ambion). Samples were incubated at 50 °C for 45 min and then boiled for 10 min. RNA was isolated using 500 µl of TRIzol reagent and the miRNeasy kit (Qiagen) with on-column DNase digestion (Qiagen). RNA was eluted with 10 µl of water and then analyzed by quantitative RT-PCR for the detection of enriched transcripts. For protein elution, beads were resuspended in three times the original volume of DNase buffer (100 mM NaCl, 0.1% NP-40), and protein was eluted with a cocktail of 100 µg/ml RNase A (Sigma-Aldrich), 0.1 U/ml RNase H (Epicenter) and 100 U/ml DNase I (Invitrogen) at 37 °C for 30 min. Eluted protein samples were supplemented with NuPAGE LDS Sample Buffer (Novex) and NuPAGE Sample Reducing Agent (Novex) to a final concentration of 1× each and then boiled for 10 min before SDS-PAGE protein blot analysis using an antibody to SNF5 (Millipore).

RNA-seq library preparation. Total RNA was extracted from healthy and cancer cell lines and subject tissues, and RNA quality was assessed via Agilent Bioanalyzer. Transcriptome libraries from the mRNA fractions were generated following the RNA-seq protocol (Illumina). Each sample was sequenced in a single lane with the Illumina Genome Analyzer II (with a 40- to 80-nt read length) or with the Illumina HiSeq 2000 (with a 100-nt read length) according to published protocols^{11,49}. For strand-specific library construction, we employed the dUTP method of second-strand marking as described previously⁵⁰.

Statistical analyses for experimental studies. All data are presented as means ± s.e.m. All experimental assays were performed in duplicate or triplicate. Statistical analyses shown in figures represent Fisher's exact tests or two-tailed *t* tests, as indicated. For details regarding the statistical methods employed during microarray, RNA-seq and ChIP-seq data analysis, see below.

Nomination of *SchLAP1* as an outlier using RNA-seq data. We nominated *SchLAP1* as a prostate cancer outlier as described¹¹. Briefly, a modified COPA analysis was performed on the 81 tissue samples in the cohort. Reads per kilobase per million mapped reads (RPKM) expression values were used and shifted by 1.0 to avoid division by zero. COPA analysis included the

following steps: (i) gene expression values were median centered, using the median expression value for the gene across all samples in the cohort, which sets the gene's median to zero; (ii) the median absolute deviation (MAD) was calculated for each gene, and each gene expression value was then scaled by its MAD; (iii) the 80th, 85th, 90th and 98th percentiles of the transformed expression values were calculated for each gene, the average of those four values was taken, and genes were then ranked according to this 'average percentile', which generated a list of outlier genes arranged by importance; and (iv) finally, genes showing an outlier profile in the benign samples were discarded.

LNCaP ChIP-seq data. Sequencing data from GSE14097 were downloaded from GEO. Reads from the LNCaP H3K4me3 and H3K36me3 ChIP-seq samples were mapped to human genome version hg19 using BWA 0.5.9 (ref. 51). Peak calling was performed using MACS⁵² according to published protocols⁵³. Data were visualized using the UCSC Genome Browser⁵⁴.

RWPE ChIP-seq data. Sequencing data from RWPE SNF5 ChIP-seq samples were mapped to human genome version hg19 using the BWA 0.5.9 algorithm⁵¹. Although we performed paired-end sequencing, the ChIP-seq reads were processed as single-end reads to adhere to our preexisting analysis protocol. Basic read alignment statistics are listed in **Supplementary Table 6a**. Peak calling was performed with respect to an IgG control using the MACS algorithm⁵². We bypassed the model-building step of MACS (using the '-nomodel' flag) and specified a shift size equal to half the library fragment size determined by the Agilent Bioanalyzer (using the '-shiftsize' option). For each sample, we ran the CEAS program and generated genome-wide reports⁵⁵. We retained peaks with an FDR less than 5% (peak calling statistics across multiple FDR thresholds are shown in **Supplementary Table 6b**). We then aggregated SNF5 peaks from the RWPE-*LacZ*, RWPE-*SchLAP1* isoform 1 and RWPE-*SchLAP1* isoform 2 samples using the 'union' of the genomic peak intervals. We intersected peaks with RefSeq protein-coding genes and found that 1,299 peaks occurred within 1 kb of TSSs. We counted the number of reads overlapping each of these promoter peaks across each sample using a custom Python script and used the DESeq R package⁵⁶ version 1.6.1 to compute the normalized fold change between RWPE-*LacZ* and RWPE-*SchLAP1* (both isoforms). We observed that 389 of the 1,299 promoter peaks had at least a 2-fold average decrease in SNF5 binding. This set of 389 genes was subsequently used as a gene set for GSEA (**Supplementary Table 6c**).

Microarray experiments. We performed two-color microarray gene expression profiling of 22Rv1 and LNCaP cells treated with two independent siRNAs targeting *SchLAP1* as well as control non-targeting siRNAs. These profiling experiments were run in technical triplicate for a total of 12 arrays (6 from 22Rv1 and 6 from LNCaP). Additionally, we profiled 22Rv1 and LNCaP cells treated with independent siRNAs targeting SWI/SNF component *SMARCB1* as well as control non-targeting siRNAs. These profiling experiments were run as biological duplicates for a total of four arrays (two cell lines × two independent siRNAs × one protein). Finally, we profiled RWPE cells expressing two different *SchLAP1* isoforms as well as the control *LacZ* gene. These profiling experiments were run in technical duplicate for a total of four arrays (two from RWPE-*SchLAP1* isoform 1 and two from RWPE-*SchLAP1* isoform 2).

Processing to determine ranked gene expression lists. All of the microarray data were represented as log₂ fold change between targeting versus control siRNAs. We used the CollapseDataset tool provided by the GSEA package to convert Agilent Probe IDs to gene symbols. Genes whose expression was measured by multiple probes were consolidated using the median values obtained with these probes. We then ran one-class SAM analysis from the Multi-Experiment Viewer application and ranked all genes by the difference between observed versus expected statistics. These ranked gene lists were imported to GSEA version 2.07.

***SchLAP1* siRNA knockdown microarrays.** For the 22Rv1 and LNCaP *SchLAP1* knockdown experiments, we ran the GseaPreRanked tool to discover enriched gene sets in MSigDB²² version 3.0. Lists of positively and negatively enriched concepts were interpreted manually.

SMARCB1 siRNA knockdown microarrays. For each *SMARCB1* knockdown experiment, we nominated genes that were altered by an average of at least twofold. These signatures of putative SNF5 target genes were then used to assess enrichment of *SchLAP1*-regulated genes using the GseaPreRanked tool. Additionally, we nominated genes whose expression changed by an average of twofold or greater across *SMARCB1* knockdown experiments and quantified the enrichment for *SchLAP1* target genes using GSEA.

RWPE *SchLAP1* expression microarrays. RWPE-*SchLAP1* versus RWPE-*LacZ* expression profiles were ranked using SAM analysis as described above. A total of 1,245 genes were significantly over- or underexpressed and are shown in **Supplementary Table 6d**. A q value of 0.0 in this SAM analysis signifies that no permutation generated a more significant difference between observed and expected gene expression ratios. The ranked gene expression list was used as input for the GseaPreRanked tool and compared against SNF5 ChIP-seq promoter peaks that decreased by >2-fold in RWPE cells overexpressing *SchLAP1*. Of the 389 genes in the ChIP-seq gene set, 250 were profiled by the Agilent Human Genome microarray chip and were present in the GSEA gene symbol database. The expression profile across these 250 genes is shown in **Supplementary Table 6e**.

RNA-seq data. We assembled an RNA-seq cohort from prostate cancer tissues sequenced at multiple institutions. We included data from 12 primary tumors and 5 benign tissues published in GEO ([GSE22260](#))⁵⁷, from 16 primary tumors and 3 benign tissues released in the database of Genotypes and Phenotypes (dbGAP) ([phs000310.v1.p1](#))⁵⁸ and from 17 benign, 57 primary and 14 metastatic tumors sequenced by our own institution and released in dbGAP ([phs000443.v1.p1](#)). Sample information is shown in **Supplementary Table 1a**, and sequencing library information is shown in **Supplementary Table 1b**.

RNA-seq alignment and gene expression quantification. Sequencing data were aligned using TopHat⁵⁹ version 1.3.1 against the Ensembl GRCh37 human genome build. Known introns (Ensembl release 63) were provided to TopHat. Gene expression across genes in Ensembl version 63 and the *SchLAP1* transcript was quantified by HT-Seq version 0.5.3p3 using the script 'htseq-count'. Reads were counted without respect to strand to avoid bias between unstranded and strand-specific library preparation methods. This bias results from the inability to resolve reads in regions where two genes on opposite strands overlap in the genome.

RNA-seq differential expression analysis. Differential expression analysis was performed using R package DESeq⁵⁶ version 1.6.1. Read counts were normalized using the 'estimateSizeFactors' function, and variance was modeled by the 'estimateDispersions' function. Statistics on differential expression were computed by the 'nbinomTest' function. We called differentially expressed genes by imposing adjusted P -value cutoffs for cancer versus benign samples ($P_{\text{adj}} < 0.05$), metastasis versus primary samples ($P_{\text{adj}} < 0.05$) and Gleason score of 8+ versus 6 ($P_{\text{adj}} < 0.10$). Heatmap visualizations of these analyses are presented as **Supplementary Figure 5**.

RNA-seq correlation analysis. Read count data were normalized using functions from the R package DESeq version 1.6.1. Adjustments for library size were made using the 'estimateSizeFactors' function, and variance was modeled using the 'estimateDispersions' function using the parameters 'method=blind' and 'sharingMode=fit-only'. Next, raw read count data were converted to pseudocounts using the 'getVarianceStabilizedData' function. Gene expression levels were then mean centered and standardized using the 'scale' function in R. Pearson's correlation coefficients were computed between each gene of interest and all other genes. Statistical significance of Pearson's correlations was determined by comparison to correlation coefficients achieved with 1,000 random permutations of the expression data. We controlled for multiple-hypothesis testing using the 'qvalue' package in R. The *SchLAP1* correlation signature of 253 genes was determined by imposing a cutoff of $q < 0.05$.

Oncomine concepts analysis of the *SchLAP1* signature. We separated the 253 genes with expression levels significantly correlated with *SchLAP1* into positively and negatively correlated gene lists. We imported these gene

lists into Oncomine as custom concepts. We then nominated significantly associated prostate cancer concepts with odds ratio > 3.0 and $P < 1 \times 10^{-6}$. We exported these results as the nodes and edges of a concept association network and visualized the network using Cytoscape version 2.8.2. Node positions were computed using the Force-Directed Layout algorithm in Cytoscape using the odds ratio as the edge weight. Node positions were subtly altered manually to enable better visualization of node labels.

Association of correlation signatures with Oncomine concepts. We applied our RNA-seq correlation analysis procedure to the genes *SchLAP1*, *EZH2*, *PCA3*, *AMACR* and *ACTB*. For each gene, we created signatures from the top 5% of positively and negatively correlated genes (**Supplementary Table 3**). We performed a large meta-analysis of these correlation signatures across Oncomine data sets corresponding to disease outcome (Glinisky Prostate and Setlur Prostate), metastatic disease (Holzbeierlein Prostate, Lapointe Prostate, LaTulippe Prostate, Taylor Prostate 3, Vanaja Prostate, Varambally Prostate and Yu Prostate), advanced Gleason score (Bittner Prostate, Glinisky Prostate, Lapointe Prostate, LaTulippe Prostate, Setlur Prostate, Taylor Prostate 3 and Yu Prostate) and localized cancer (Arredouani Prostate, Holzbeierlein Prostate, Lapointe Prostate, LaTulippe Prostate, Taylor Prostate 3, Varambally Prostate and Yu Prostate). We also incorporated our own concept signatures for metastasis, advanced Gleason score and localized cancer determined from our RNA-seq data. For each concept, we downloaded the gene signatures corresponding to the top 5% of genes up- and downregulated. Pairwise signature comparisons were performed using a one-sided Fisher's exact test. We controlled for multiple-hypothesis testing using the 'qvalue' package in R. We considered concept pairs with $q < 0.01$ and odds ratio > 2.0 as significant. In cases where a gene signature associated with both the over- and underexpression gene sets from a single concept, only the most significant result (as determined by odds ratio) is shown.

Analysis of *SchLAP1* and *SMARCB1* expression signatures. Gene signatures obtained with knockdown of *SchLAP1* and *SMARCB1* were generated from Agilent gene expression microarray data sets. For each cell line, we obtained a single vector of per-gene fold changes by averaging technical replicates and then taking the median across biological replicates. We merged the results from individual cell line using the median of the changes in 22Rv1 and LNCaP cells. Venn diagram plots were produced using the BioVenn website⁶⁰. We then compared the top 10% of upregulated and downregulated genes with knockdown of *SchLAP1* and *SMARCB1* to gene signatures downloaded from the Taylor Prostate 3 data set in the Oncomine database. We performed signature comparison using one-sided Fisher's exact tests and controlled for multiple testing using the R package 'qvalue'. Signature comparisons with $q < 0.05$ were considered significantly enriched. We plotted the odds ratios from significant comparisons using the 'heatmap.2' function in the 'gplots' R package.

Kaplan-Meier survival analysis based on the *SchLAP1* gene signature. We downloaded prostate cancer expression profiling data and clinical annotations from [GSE8402](#), published by Setlur *et al.*¹⁷. We intersected the 253-gene *SchLAP1* signature with the genes in this data set and found 80 genes in common. We then assigned *SchLAP1* expression scores to each patient sample in the cohort using the unweighted sum of standardized expression levels across the 80 genes. Given that we observed *SchLAP1* expression in approximately 20% of prostate cancer samples, we used the 80th percentile of *SchLAP1* expression scores as the threshold for 'high' versus 'low' scores. We then performed 10-year survival analysis using the 'survival' package in R and computed statistical significance using the log-rank test.

Additionally, we imported the 253-gene *SchLAP1* signature into Oncomine to download the expression data for 167 of the 253 genes profiled by the Glinisky prostate data set¹⁶. We assigned *SchLAP1* expression scores in a similar fashion and designated the top 20% of patients as having 'high' *SchLAP1* scores. We performed survival analysis using the time to biochemical prostate-specific antigen (PSA) recurrence and computed statistical significance as described above.

PhyloCSF analysis. We obtained 46-way multi-alignment FASTA files for *SchLAP1*, *HOTAIR*, *GAPDH* and *ACTB* using the 'Stitch Gene blocks' tool

within the Galaxy bioinformatics framework. We evaluated each gene for the likelihood that it represented a protein-coding region using PhyloCSF software (version released 28 October 2012). Each gene was evaluated using the phylogeny from 29 mammals (available by default within PhyloCSF) in any of the 3 reading frames. Scores are measured in decibans and represent the likelihood ratio that a sequence is protein-coding rather than noncoding.

Mayo Clinic cohort analyses. Subjects were selected from a cohort of individuals from the Mayo Clinic with high-risk prostate cancer who had undergone radical prostatectomy. The cohort was defined as 1,010 men with high-risk prostate cancer who underwent radical prostatectomy between 2000 and 2006, of whom 73 developed clinical progression (defined as individuals with systemic disease as evidenced by positive bone or computed tomography (CT) scan)⁶¹. High risk of recurrence was defined by preoperative PSA levels of >20 ng/ml, pathological Gleason score of 8–10, seminal vesicle invasion (SVI) or Gleason, PSA, seminal vesicle and margin (GPSM) score of ≥ 10 (ref. 62). The subcohort incorporated all 73 subjects with clinical progression to systemic disease and a random sampling of 20% of the entire cohort (202 men, including 19 with clinical progression). The total case-cohort study included 256 subjects, and tissue specimens were available from 235 subjects. The subcohort was previously used to validate a genomic classifier for predicting clinical progression⁶¹.

Tissue preparation. Formalin-fixed, paraffin-embedded samples of human prostate adenocarcinoma prostatectomies were collected from subjects with informed consent at the Mayo Clinic according to an IRB-approved protocol. Pathological review of tissue sections stained with hematoxylin and eosin was used to guide macrodissection of the tumor from surrounding stromal tissue in three to four 10- μ m sections. The index lesion was considered as the dominant lesion by size.

RNA extraction and microarray hybridization. For the validation cohort, total RNA was extracted and purified using a modified protocol for the commercially available RNeasy FFPE nucleic acid extraction kit (Qiagen). RNA concentrations were calculated using a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies). Purified total RNA was subjected to whole-transcriptome amplification using the WT-Ovation FFPE system according to the manufacturer's recommendation with minor modifications (NuGen). For the validation, only the Ovation FFPE WTA System was used. Amplified products were fragmented and labeled using the Encore Biotin Module (NuGen) and hybridized to Affymetrix Human Exon (HuEx) 1.0 ST GeneChips following the manufacturer's recommendations.

Microarray expression analysis. Normalization and summarization of the microarray samples was performed with the frozen Robust Multiarray Average (fRMA) algorithm using custom frozen vectors. These custom vectors were created using the vector creation methods described previously⁶³. Quantile normalization and robust weighted average methods were used for normalization and summarization, respectively, as implemented in fRMA.

Statistical analysis. Given the exon-intron structure of isoform 1 of *SChLAP1*, all probe selection regions (or PSRs) that fell within the genomic span of *SChLAP1* were inspected for overlap with any of the exons of this gene. One PSR, 2518129, was found to be fully nested within exon 3 of *SChLAP1*

and was used for further analysis as a representative PSR for this gene. The PAM (Partition Around Medoids) unsupervised clustering method was used on the expression values of all clinical samples to define two groups with high and low expression of *SChLAP1*.

Statistical analysis on the association of *SChLAP1* with clinical outcomes was carried out using three endpoints: (i) biochemical recurrence, defined as two consecutive increases in serum PSA of ≥ 0.2 ng/ml after radical prostatectomy; (ii) clinical progression, defined as a positive CT or bone scan; and (iii) prostate cancer-specific mortality.

For the clinical progression end point, all subjects with clinical progression were included in the survival analysis, whereas controls in the subcohort were weighted in a fivefold manner to be representative of individuals from the original cohort. For the prostate cancer-specific mortality end point, cases who did not die from prostate cancer were omitted, and weighting was applied in a similar manner. For biochemical recurrence, because the case cohort was designed on the basis of the clinical progression end point, resampling of subjects with biochemical recurrence and the subcohort was performed to have a representative of the selected individuals with biochemical recurrence from the original cohort.

46. Rubin, M.A. *et al.* Rapid ("warm") autopsy study for procurement of metastatic prostate cancer. *Clin. Cancer Res.* **6**, 1038–1045 (2000).
47. Tomlins, S.A. *et al.* Role of the *TMPRSS2-ERG* gene fusion in prostate cancer. *Neoplasia* **10**, 177–188 (2008).
48. Chu, C., Qu, K., Zhong, F.L., Artandi, S.E. & Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
49. Maher, C.A. *et al.* Chimeric transcript discovery by paired-end transcriptome sequencing. *Proc. Natl. Acad. Sci. USA* **106**, 12353–12358 (2009).
50. Levin, J.Z. *et al.* Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat. Methods* **7**, 709–715 (2010).
51. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
52. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
53. Feng, J., Liu, T. & Zhang, Y. Using MACS to identify peaks from ChIP-Seq data. *Curr. Protoc. Bioinformatics* **Chapter 2** Unit 2.14 (2011).
54. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
55. Shin, H., Liu, T., Manrai, A.K. & Liu, X.S. CEAS: *cis*-regulatory element annotation system. *Bioinformatics* **25**, 2605–2606 (2009).
56. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
57. Kannan, K. *et al.* Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proc. Natl. Acad. Sci. USA* **108**, 9172–9177 (2011).
58. Pflueger, D. *et al.* Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. *Genome Res.* **21**, 56–67 (2011).
59. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
60. Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn—a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488 (2008).
61. Karnes, R.J. *et al.* Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk patient population. *J. Urol.* doi:10.1016/j.juro.2013.06.017 (11 June 2013).
62. Blute, M.L., Bergstralh, E.J., Iocca, A., Scherer, B. & Zincke, H. Use of Gleason score, prostate specific antigen, seminal vesicle and margin status to predict biochemical failure after radical prostatectomy. *J. Urol.* **165**, 119–125 (2001).
63. Vergara, I.A. *et al.* Genomic "dark matter" in prostate cancer: exploring the clinical utility of ncRNA as biomarkers. *Front. Genet.* **3**, 23 (2012).

The lncRNAs PCGEM1 and PRNCR1 are not implicated in castration resistant prostate cancer

John R. Prensner^{1,*}, Anirban Sahu^{1,*}, Matthew K. Iyer^{1,2,*}, Rohit Malik^{1,*}, Benjamin Chandler¹, Irfan A. Asangani¹, Anton Poliakov¹, Ismael A. Vergara³, Mohammed Alshalalfa³, Robert B. Jenkins⁴, Elai Davicioni³, Felix Y. Feng^{1,5,7}, Arul M. Chinnaiyan^{1,2,6,7,8}

¹ Michigan Center for Translational Pathology, University of Michigan, Ann Arbor, Michigan USA.

² Department of Computational Medicine and Bioinformatics, Ann Arbor, Michigan USA.

³ GenomeDx Biosciences Inc., Vancouver, British Columbia, Canada.

⁴ Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota USA.

⁵ Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan USA.

⁶ Department of Pathology, University of Michigan, Ann Arbor, Michigan USA.

⁷ Comprehensive Cancer Center, University of Michigan, Ann Arbor, Michigan USA.

⁸ Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan USA.

* These authors contributed equally

Correspondence to: Arul M. Chinnaiyan, **email:** arul@med.umich.edu

Keywords: prostate cancer, long noncoding RNA, androgen receptor

Received: February 5, 2014

Accepted: March 21, 2014

Published: March 23, 2014

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT:

Long noncoding RNAs (lncRNAs) are increasingly implicated in cancer biology, contributing to essential cancer cell functions such as proliferation, invasion, and metastasis. In prostate cancer, several lncRNAs have been nominated as critical actors in disease pathogenesis. Among these, expression of *PCGEM1* and *PRNCR1* has been identified as a possible component in disease progression through the coordination of androgen receptor (AR) signaling (Yang et al., Nature 2013, see ref. [1]). However, concerns regarding the robustness of these findings have been suggested. Here, we sought to evaluate whether *PCGEM1* and *PRNCR1* are associated with prostate cancer. Through a comprehensive analysis of RNA-sequencing data (RNA-seq), we find evidence that *PCGEM1* but not *PRNCR1* is associated with prostate cancer. We employ a large cohort of >230 high-risk prostate cancer patients with long-term outcomes data to show that, in contrast to prior reports, neither gene is associated with poor patient outcomes. We further observe no evidence that *PCGEM1* nor *PRNCR1* interact with AR, and neither gene is a component of AR signaling. Thus, we conclusively demonstrate that *PCGEM1* and *PRNCR1* are not prognostic lncRNAs in prostate cancer and we refute suggestions that these lncRNAs interact in AR signaling.

INTRODUCTION

Long noncoding RNAs (lncRNAs) have emerged as a critical element in cell biology, contributing to a wide variety of cellular behaviors and functions [2]. In cancer, lncRNAs have been the subject of much research during the past five years. Notably, lncRNAs are known to coordinate aggressive phenotypes of several common

tumors, including breast cancer and prostate cancer [3, 4]. Large profiling studies have suggested that upwards of 10,000 lncRNAs may exist in the human genome [5]; yet only a fraction of these entities have been characterized. Thus, the identity and function of lncRNAs in cancer remains largely unknown.

In prostate cancer, several lncRNAs, including *PCA3* and *PCAT-1*, have been shown to be upregulated

in patients with cancer [6-9]. Recently, two lncRNAs, *PCGEM1* and *PRNCR1*, have been suggested in prostate cancer to act as mediators of castration-resistance disease by binding, in a direct and sequential fashion, to the androgen receptor (AR), causing ligand-independent activation of its gene expression programs [1]. While *PCGEM1* has been observed in prostate cancer previously [6, 10], *PRNCR1* is a poorly characterized transcript, and we were concerned that *PRNCR1* had not been nominated by previous global profiling studies of prostate cancers [7, 11-14].

We therefore sought to investigate *PRNCR1* and *PCGEM1* in prostate cancer. In specific, we sought to reproduce three core observations suggested by Yang et al published in *Nature* (see [1]) and include: 1) that *PRNCR1* and *PCGEM1* are highly overexpressed in aggressive forms of prostate cancer, 2) that these two lncRNAs bind to AR under ligand-stimulated conditions, and 3) that the coordination of *PRNCR1* and *PCGEM1* interact with AR via specific post-translational modifications of the AR protein. Here, we report that none of these three findings is fully reproducible.

First, we asked whether *PCGEM1* and *PRNCR1* are highly overexpressed in aggressive prostate cancer, as suggested by others (see [1, 15]). Indeed, while some have argued that these lncRNAs are critical in castration-resistant prostate cancer [1], there has been no study that evaluated the expression of these lncRNAs in tissue samples from human castrate-resistant prostate cancers (CRPC). To evaluate these lncRNAs in more detail, we first assessed their expression levels in 171 human prostatic tissues using RNA sequencing data aggregated from four independent studies of prostate cancer, including our own internal datasets [1, 12-14] (Fig. 1A). Whereas we found robust expression of *PCGEM1* in a subset of prostate tissues (RPKM >1 in 82 samples; RPKM >10 in 27 samples), we observed scant levels of *PRNCR1* in all samples (RPKM >1 in only 3 samples; RPKM >10 in 0 samples) (Supplementary Table 1). This does not lend confidence to *PRNCR1* as a significant entity in this disease. For comparison, we used the prostate cancer lncRNA *SchLAP1* as a positive control. We found extreme overexpression of *SchLAP1* in samples from all datasets (RPKM >1 in 69 samples; RPKM >10 in 26 samples) (Supplementary Fig. 1). To rule out the possibility that *PRNCR1* was a non-poly-adenylated RNA, we verified experimentally that *PRNCR1* was observed in the poly-A fraction of RNA that was used to generate the RNA-seq data (Supplementary Fig. 2).

Given the low support for *PRNCR1* in the RNA-seq data, we next confirmed these findings using quantitative PCR (qPCR) in a large set of prostate cancer tissues including 34 PCAs and 31 CRPC tumors as well as 18 benign adjacent tissues. As shown in Fig. 1B, *PCGEM1* is upregulated in clinically localized cancer, confirming the known literature [6, 10]; however *PRNCR1* expression

does not demonstrate a convincing association with prostate cancer. We found a borderline decrease in *PRNCR1* expression in metastatic castrate-resistant cancer ($p = 0.047$, Student's t-test). We used *PCAT1*, *EZH2*, and *SchLAP1* as control genes, all of which have elevated expression in prostate cancer metastases. Conversely, we used *PCA3* as a control gene that is known to be upregulated in localized prostate cancer but not metastatic prostate cancer. Finally, while *PCGEM1* is upregulated in cancer patients from matched tumor/benign samples, *PRNCR1* does not convincingly exhibit this pattern of upregulation (Supplementary Fig. 3).

Next, an independent analysis of 235 high-risk prostate cancer tissues demonstrated that neither *PCGEM1* nor *PRNCR1* is associated with aggressive prostate cancer, and neither lncRNA stratifies prostate cancer-specific mortality (Fig. 1C,D and Supplementary Tables 2,3). An analysis of intermediate endpoints such as biochemical recurrence and progression to metastatic disease demonstrated a trend for *PCGEM1* and *PRNCR1* to be associated with less aggressive disease and favorable outcomes (Supplementary Fig. 4), which contradicts previous claims that these lncRNAs are involved in an aggressive patient clinical course [1, 15]. Using an independent validation cohort of tissues we verified that neither *PCGEM1* nor *PRNCR1* is associated with aggressive prostate cancer (Supplementary Table 2). By contrast, we have used these datasets to confirm the prognostic utility of the lncRNA *SchLAP1* in prostate cancer, and high expression of *SchLAP1* is a powerful predictor for poor patient survival (Fig. 1E) [4].

Next, we examined whether *PCGEM1* and *PRNCR1* interacted with AR. We performed RNA-IP (RIP) assays using two independent AR antibodies, including the same antibody that was previously used to show an interaction between these lncRNAs and AR [1]. In accordance with the published literature, we performed a time-series of RIP experiments following AR stimulation, because prior data suggests that these lncRNAs bind AR from 1-2 hours after AR stimulation but not at 4 hours post-stimulation [1]. In our RIP experiments, we could not confirm that AR binds to *PCGEM1* or *PRNCR1* at either 1 hour or 4 hours post-stimulation with DHT (Fig. 1F and Supplementary Fig. 5). Similarly, in cells grown at steady-state, we used a second AR antibody and did not observe binding between AR and *PCGEM1* or *PRNCR1* (Supplementary Fig. 6). DHT-stimulated cells also demonstrated no induction in *PCGEM1* or *PRNCR1* expression (Supplementary Fig. 7). These results imply that *PCGEM1* and *PRNCR1* are not AR-interacting lncRNAs.

Finally, earlier data propose that *PCGEM1* and *PRNCR1* interact with AR via specific post-translational modifications (PTMs), specifically K349 methylation (K349Me) for *PCGEM1* and K631/K634 acetylation (K631Ac/K634Ac) for *PRNCR1* [1]. To search for these PTMs, we independently performed mass spectrometry for

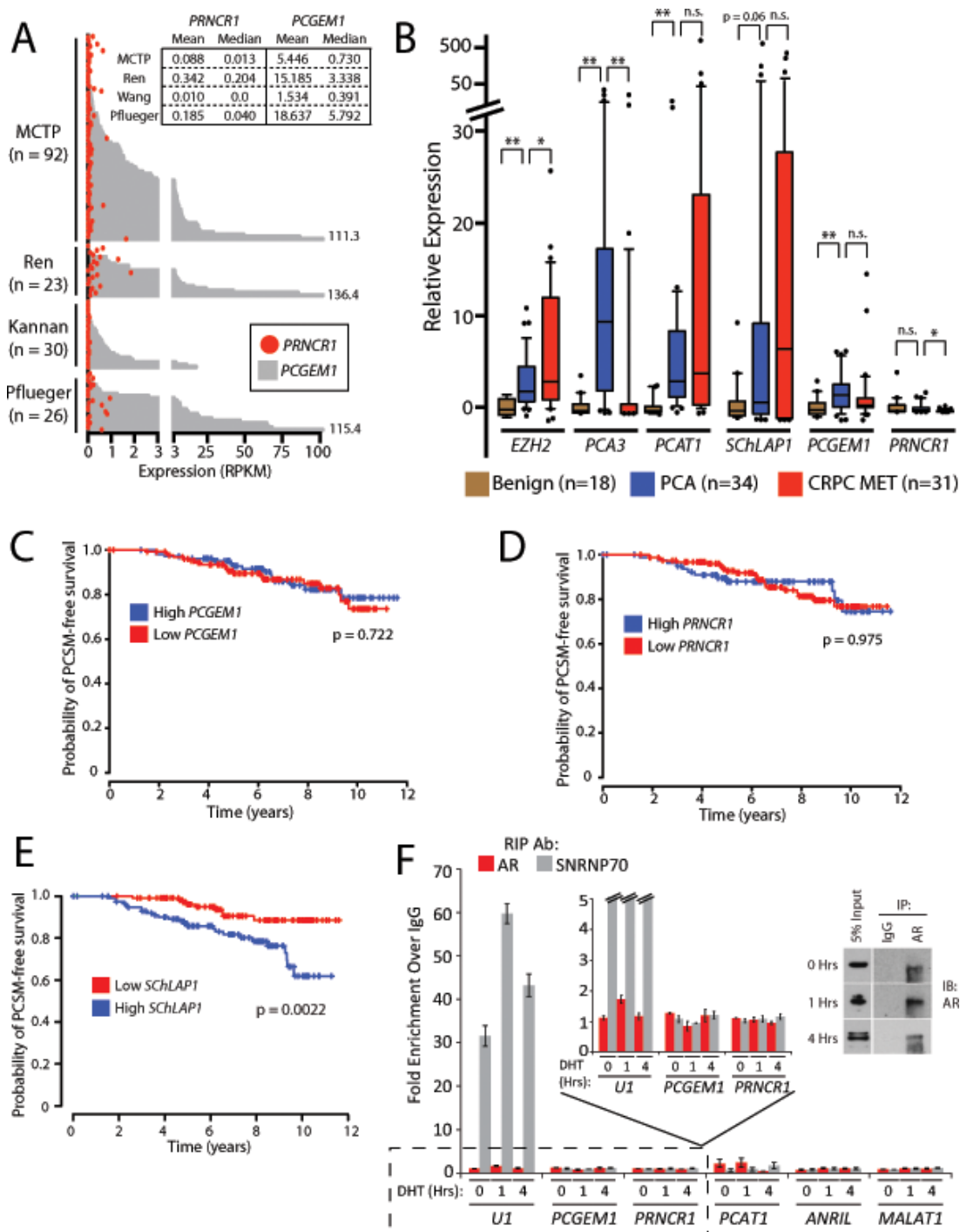


Figure 1: PCGEM1 and PRNCR1 are not associated with prostate cancer progression and do not bind the androgen receptor. (A) Plot showing *PCGEM1* (grey bars) and *PRNCR1* (red circles) expression levels (Reads per Kilobase per Million Reads, or RPKM) across 171 samples from four RNA-Seq studies of prostate cancer: Michigan Center for Translational Pathology (MCTP, internal data and dbGAP, phs000443.v1.p1), Ren et al. [13] (EGA, ERP00550), Kannan et al. [14] (GEO, GSE22260), and Pflueger et al. [12] (dbGAP, phs000310.v1.p1). Inset box shows descriptive statistics for each study. (B) Quantitative PCR for *PCGEM1* and *PRNCR1* in a cohort of prostate cancer tissues, benign (n = 18), localized cancer (n = 34), metastatic cancer (n = 31). An asterisk (*) indicates $p < 0.05$. Two asterisks (**) indicate $p < 0.01$. n.s. = non-significant. P values were determined by a two-tailed Student's t-test. Data for *SchLAP1* is obtained and re-analyzed from a prior publication (ref. [4]). (C) *PCGEM1* expression does not predict for prostate cancer-specific mortality (PCSM). (D) *PRNCR1* expression does not predict for PCSM. (E) High *SchLAP1* expression is a powerful predictor of PCSM ($p = 0.0022$). Data in (E) is reproduced from a prior publication (ref. [4]). P values in (C-E) are determined using a log-rank test. (F) RNA-immunoprecipitation (RIP) for AR following stimulation of LNCaP cells with 100nM DHT does not show binding of *PRNCR1* or *PCGEM1* to AR. *U1* binding to SNRNP70 is used as a positive control. *PCAT-1*, *ANRIL*, and *MALAT1* serve as negative controls. Inset: Western blot confirmation of AR protein pull-down by the immunoprecipitation assays. Error bars represent S.E.M.

AR in the LNCaP cell line, achieving 95% coverage of all possible tryptic peptides. We were unable to confirm that these PTMs (K349Me, K631Ac, or K634Ac) are present on AR (Supplementary Fig. 8 and Supplementary Table 4). To examine this discrepancy further, we re-analyzed prior AR MS data (found in [1]). Although this MS dataset was obtained with a trypsin digestion to prepare samples for MS, we found no fully tryptic peptides supporting the nomination of K349Me, K631Ac, or K634Ac (Supplementary Fig. 8). In fact, in the MS data for these PTMs in ref. [1], almost all peptides harboring these PTMs are non-tryptic, which are generally considered to be analysis artifacts since true non-tryptic peptides are exceedingly rare following a trypsin digestion [16-18] (Supplementary Discussion). Non-tryptic peptides are also associated with a high false-discovery rate [19]. All peptides nominating the K349Me, K631Ac, or K634Ac PTMs in ref. [1] also had multiple additional PTMs that were nominated, indicating non-specificity. These included extraordinarily rare and unusual PTMs such as oxidated lysine and deamidated asparagine, which suggest technical artifacts given the negligible likelihood of multiple rare and unusual PTMs occurring on true non-tryptic peptides. The statistical confidence for these non-tryptic peptides is <5%, whereas the corresponding fully tryptic peptides for these amino acid residues had statistical confidences >90%.

In summary, we have been unable to show a convincing role for *PCGEM1* or *PRNCRI* in aggressive prostate cancer or AR signaling. First, our data analysis of numerous human prostate cancer tissues from multiple independent laboratories indicates that neither *PCGEM1* nor *PRNCRI* are associated with castration-resistant prostate cancer. Second, we were unable to verify that *PCGEM1* and *PRNCRI* bind to the androgen receptor. Lastly, we are unconvinced that the K349Me, K631Ac, or K634Ac AR PTMs represent a plausible mechanism for interaction between AR and *PCGEM1* and *PRNCRI*. While our results challenge the notion that *PCGEM1* and *PRNCRI* play a causal role in prostate cancer, we regard lncRNAs as an emerging field of study in cancer [3, 6, 20, 21] and we are encouraged by the interest in lncRNAs in prostate cancer.

METHODS

Prostate tissues were obtained from the radical prostatectomy series and Rapid Autopsy Program at the University of Michigan tissue core. All tissue samples were collected with informed consent under an Institutional Review Board (IRB) approved protocol at the University of Michigan. Outcomes analyses were performed on a cohort of Mayo Clinic prostate cancer radical prostatectomy samples obtained under an IRB-approved protocol as described previously. Cell lines were maintained according to standard conditions. For

RIP experiments, cells were deprived of androgen for 48 hours prior to stimulation with 100nM DHT. RIP experiments were performed as previously described [1, 4]. Bioinformatics analyses utilized publicly available RNA-Seq data. Please see Supplementary Methods for details.

ACKNOWLEDGEMENTS

We thank Xia Jiang and Shruthi Subramaniam for technical assistance. This work was supported in part by the NIH Prostate Specialized Program of Research Excellence grant P50CA69568, the Early Detection Research Network grant UO1 CA111275, the US National Institutes of Health R01CA132874-01A1, and the Department of Defense grant PC100171 (A.M.C.). A.M.C. is supported by the Doris Duke Charitable Foundation Clinical Scientist Award, the Prostate Cancer Foundation, and the Howard Hughes Medical Institute. A.M.C. is an American Cancer Society Research Professor and a Taubman Scholar of the University of Michigan. F.Y.F. was supported by the Prostate Cancer Foundation, the Department of Defense grant PC094231. J.R.P. was supported by a Prostate Cancer Foundation Young Investigator Award. A.S. was supported by the NIH Predoctoral Fellowship 1F30CA180376-01. M.K.I. was supported by the Department of Defense Predoctoral Fellowship BC100238. R.M. was supported by the Department of Defense Post-doctoral Fellowship W81XWH-13-1-0284. J.R.P., M.K.I., and A.S. are Fellows of the University of Michigan Medical Scientist Training Program.

Disclosures and Competing Financial Interests

The University of Michigan has filed a patent on lncRNAs in prostate cancer, including *SchLAPI*, in which A.M.C., J.R.P. and M.K.I. are named as co-inventors. Wafergen, Inc. has a non-exclusive license for creating commercial research assays for lncRNAs in prostate cancer. GenomeDx Biosciences Inc has an exclusive license for creating tissue assays for lncRNAs in prostate cancer. A.M.C. is a co-founder and advisor to Compendia Biosciences, which supports the Oncomine database. He also serves on the Scientific Advisory Board of Wafergen; neither Life Technologies or Wafergen had any role in the design or experimentation of this study, nor have they participated in the writing of the manuscript. I.A.V. and E.D. are employees of GenomeDx Biosciences Inc.

Author Contributions

J.R.P., R.M., M.K.I., A.S. and A.M.C. designed the project and directed experimental studies. J.R.P., R.M., A.S. and B.C. performed *in vitro* studies. M.K.I. performed

bioinformatics analysis. I.A.A. and A.P. performed AR mass spectrometry. I.A.V., R.B.J., E.D., and M.A. acquired human tissue samples and performed statistical outcomes analyses for *PCGEM1* and *PRNCR1* expression. J.R.P., M.K.I., A.S., R.M., F.Y.F. and A.M.C. designed experiments, interpreted data, and wrote the manuscript.

REFERENCES

1. Yang L, Lin C, Jin C, Yang JC, Tanasa B, Li W, Merkurjev D, Ohgi KA, Meng D, Zhang J, Evans CP and Rosenfeld MG. lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature*. 2013; 500(7464):598-602.
2. Rinn JL and Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem*. 2012; 81:145-166.
3. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*. 2010; 464(7291):1071-1076.
4. Prensner JR, Iyer MK, Sahu A, Asangani IA, Cao Q, Patel L, Vergara IA, Davicioni E, Erho N, Ghadessi M, Jenkins RB, Triche TJ, Malik R, Bedenis R, McGregor N, Ma T, et al. The long noncoding RNA SCHLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nat Genet*. 2013; 45(11):1392-1398.
5. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A and Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev*. 2011; 25(18):1915-1927.
6. Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, Chen Y and Liu XS. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat Struct Mol Biol*. 2013; 20(7):908-913.
7. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, Laxman B, Asangani IA, Grasso CS, Kominsky HD, Cao X, Jing X, Wang X, Siddiqui J, Wei JT, Robinson D, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. *Nat Biotechnol*. 2011; 29(8):742-749.
8. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeny LA, Aalders TW, Swinkels DW and Schalken JA. DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res*. 2002; 62(9):2695-2698.
9. Hessels D and Schalken JA. The use of PCA3 in the diagnosis of prostate cancer. *Nat Rev Urol*. 2009; 6(5):255-261.
10. Srikantan V, Zou Z, Petrovics G, Xu L, Augustus M, Davis L, Livezey JR, Connell T, Sesterhenn IA, Yoshino K, Buzard GS, Mostofi FK, McLeod DG, Moul JW and Srivastava S. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. *Proc Natl Acad Sci U S A*. 2000; 97(22):12216-12221.
11. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS, Arora VK, Kaushik P, Cerami E, Reva B, Antipin Y, Mitsiades N, Landers T, Dolgalev I, Major JE, Wilson M, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010; 18(1):11-22.
12. Pflueger D, Terry S, Sboner A, Habegger L, Esgueva R, Lin PC, Svensson MA, Kitabayashi N, Moss BJ, MacDonald TY, Cao X, Barrette T, Tewari AK, Chee MS, Chinnaiyan AM, Rickman DS, et al. Discovery of non-ETS gene fusions in human prostate cancer using next-generation RNA sequencing. *Genome Res*. 2011; 21(1):56-67.
13. Ren S, Peng Z, Mao JH, Yu Y, Yin C, Gao X, Cui Z, Zhang J, Yi K, Xu W, Chen C, Wang F, Guo X, Lu J, Yang J, Wei M, et al. RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. *Cell Res*. 2012; 22(5):806-821.
14. Kannan K, Wang L, Wang J, Ittmann MM, Li W and Yen L. Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. *Proc Natl Acad Sci U S A*. 2011; 108(22):9172-9177.
15. Petrovics G, Zhang W, Makarem M, Street JP, Connelly R, Sun L, Sesterhenn IA, Srikantan V, Moul JW and Srivastava S. Elevated expression of PCGEM1, a prostate-specific gene with cell growth-promoting function, is associated with high-risk prostate cancer patients. *Oncogene*. 2004; 23(2):605-611.
16. Shilov IV, Seymour SL, Patel AA, Loboda A, Tang WH, Keating SP, Hunter CL, Nuwaysir LM and Schaeffer DA. The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics*. 2007; 6(9):1638-1655.
17. Kim JS, Monroe ME, Camp DG, 2nd, Smith RD and Qian WJ. In-source fragmentation and the sources of partially tryptic peptides in shotgun proteomics. *J Proteome Res*. 2013; 12(2):910-916.
18. Picotti P, Aebersold R and Domon B. The implications of proteolytic background for shotgun proteomics. *Mol Cell Proteomics*. 2007; 6(9):1589-1598.
19. Olsen JV, Ong SE and Mann M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics*. 2004; 3(6):608-614.
20. Prensner JR and Chinnaiyan AM. The emergence of lncRNAs in cancer biology. *Cancer Discov*. 2011; 1(5):391-407.
21. Kretz M, Siprashvili Z, Chu C, Webster DE, Zehnder A, Qu K, Lee CS, Flockhart RJ, Groff AF, Chow J, Johnston D, Kim GE, Spitale RC, Flynn RA, Zheng GX, Aiyer S, et al. Control of somatic tissue differentiation by the long non-coding RNA TINCR. *Nature*. 2013; 493(7431):231-235.

ROHIT MALIK, Ph.D.
Research Fellow

1400 E Medical Center Drive
Room No. 7431, Comprehensive Cancer Center
University of Michigan, Ann Arbor, MI-48105
E-mail romalik@med.umich.edu

Education and Training

July 2000-July 2003	BS in Microbiology, Panjab University, Chandigarh, India,
August 2003-July 2005	MS in Microbiology, Panjab University, Chandigarh, India,
February 2006	Post Graduate Diploma in Bioinformatics (Online), Bioinformatics Institute of India,
July-2006-December 2011	Ph.D. in Molecular Biology, Loyola University Chicago, IL, (with Distinction)
January 2012 – September 2013	Research Fellow, Michigan Center for Translational Pathology, University of Michigan

Research Experience

June 2004 – August 2005	Graduate Student (MS) at Panjab University, India
September 2005 – June 2006	DNA Analyst at LabIndia Instruments Pvt. Ltd., India
April 2007 – October 2007	Graduate Student (PhD) at the Loyola University Chicago
November 2007 – December 2011	Graduate Student (PhD) at the Loyola University Chicago
January 2012 – September 2013	Research Fellow at the University of Michigan
October 2013 – Present	Research investigator at the University of Michigan

Research Interest

- Discovery and characterization of novel long non-coding RNA associated with prostate cancer
- To evaluate lncRNAs as prognostic and diagnostic biomarkers.
- Identification and therapeutic targeting of novel androgen receptor co-activators in lethal castration resistant prostate cancer.

Grants/Research Support

Present and Active

351884	
Prostate Cancer Foundation Young Investigator Award	01/01/14-12/31/16
Investigator	\$75,000/yr
Characterization and Therapeutic Targeting of Androgen Receptor Co-activators in Castration Resistant Prostate Cancer	

W81XWH-13-1-0284

Department of Defense Post-doctoral Fellowship

08/01/13-07/31/15

Investigator

\$61,477/yr

Biological and Clinical Characterization of Novel lncRNAs Associated with Metastatic Prostate Cancer

Completed

0910098G

American Heart Association Pre-doctoral Fellowship

01/01/09-12/31/10

Investigator

\$26,000/yr

Role of Arrestins in Down-regulation of Chemokine Receptors

Honors and Awards

July 2004	Paramhansa Yogananda foundation Scholarship for Masters research
August 2005	Student travel award, 16th Evergreen International Phage Biology Meeting
October 2009	Best poster award, Molecular Biology Program research retreat
September 2011	Best poster award, Molecular biology program research retreat
April 2011	Student travel award, Experimental Biology Meeting-2011
December 2011	Thesis Distinction award, Loyola University Chicago
April 2013	Scholar-in-Training Award, Prostate Cancer Foundation and AACR
December 2013	Team Science Award, Michigan Center for Translational Pathology

Memberships in Professional Societies

October 2011 – present American Association for Cancer Research (AACR)

Editorial Positions, Boards, and Peer-Review Service

Reviewer: International Journal of Cancer, PLoS One, European Journal of Pharmacology, Anti-Cancer Research, Cell Biology International, Cancer Cell International

November 2011 – present Member faculty search committee, Michigan Center for Translational Pathology

Teaching/ Mentoring Experience

Undergraduate Research Opportunity Program (UROP) students:

September 2012 - Present	Shruthi Subramaniyan (U of M)
September 2013 - April 2014	Sonoma Patel (U of M)
September 2012 – April 2012	Alexander Carley (U of M)

Rotation Graduate (PhD) Students: University of Michigan

May 2014 - June 2014	Yajia Jiang (U of M)
May 2013 - June 2013	Anjan Saha (U of M)

Summer Students, University of Michigan

May 2014 - June 2014	Sahr Yajdani (Skyline High School)
May 2014 - June 2014	Akash Halagur
May 2013 - June 2013	Sahr Yajdani (Skyline High School)

American Society of Pharmacology and Therapeutics (ASPET) summer research students

May 2011 - June 2011	Phil Hodur
May 2010 - June 2010	Lilly Tan

Invited Talks

February 2005	Advances in Biomedical Sciences, Shivalik Institute of Paramedical Technology.
October 2010	ASBMB “Biochemistry and Cell Biology of ESCRTs in Health and Disease”.
April 2011	Experimental Biology Session: New Roles for Arrestins in Signaling, Trafficking.
March 2013	Sixth Annual Prostate Cancer Program Retreat
April 2013	AACR 2013, Session: Noncoding RNA Regulation in Cancer
March 2014	Seventh Annual Prostate Cancer Program Retreat

Patents

Rohit Malik and Adriano Marchese. The arrestin-2/STAM-1 complex as a therapeutic target to treat cancer metastasis. *United States Patent Application 20120059044*

Bibliography

Submitted/ in-revision

1. **Rohit Malik**, Amjad P. Khan, John R. Prensner, Xiaoju Wang, Matthew K. Iyer, Yang Shi, Xia Jiang, June Escara-Wilke, Rachell Brendenis, Dmitry Borkin, Anastasia K. Yocum, Dattatreya Mellacheruvu, Yuanyuan Qiao, Irfan Asangani, Yi-Mi Wu, Xuhong Cao, Felix Y. Feng, Jolanta Grembecka, Tomasz Cierpicki, Arul M. Chinnaiyan. *Targeting the MLL complex in Castration Resistant Prostate Cancer (Submitted)*.
2. O. Alejandro Balbin, **Rohit Malik**, Saravana M. Dhanasekaran, John R. Prensner, Xuhong Cao, Yi-Mi Wu, Dan Robinson, Rui Wang, Guoan Chen, David G. Beer, Alexey I. Nesvizhskii and Arul M. Chinnaiyan. *The landscape of antisense gene expression in human cancers. (Submitted)*
3. Matthew K. Iyer, Yashar S. Niknafs, Terrence R. Barrette, Anirban Sahu, **Rohit Malik**, Yasuyuki Hosono, Joseph R. Evans, John R. Prensner, Xuhong Cao, Saravana M. Dhanasekaran, Yi-Mi Wu, Dan R. Robinson, David G. Beer, Felix Y. Feng, Hariharan K. Iyer, Arul M. Chinnaiyan. *The Landscape of Long Non-Coding RNAs in Cancer. (Submitted)*
4. Sunita Shankar, **Rohit Malik**, Vishalakshi Krishnan, Shanker Kalyana-Sundaram, Anastasia Yocum, Anton Poliakov, Vishal Kothari, Xiaojun Jing, Harika Gundlapalli, Xuhong Cao, Xiaoju Wang,

Saravana M. Dhanasekaran, Nils Walter, Chandan Kumar-Sinha, Arul M. Chinnaiyan. *RAS Engagement of AGO2 Attenuates RNA Silencing to Promote Oncogenesis*. **(In revision)**

5. Saravana M. Dhanasekaran, O. Alejandro Balbin, Guoan Chen, Ernest Nadal, Shanker Kalyana-Sundaram, Jincheng Pan, Brendan Veeneman, **Rohit Malik**, Xuhong Cao, Rui Wang, Stephanie Huang, Jinjie Zhong, Xiaojun Jing, Pankaj Vats, Matthew Iyer, Yi-Mi Wu, Paul W. Harms, Jules Lin, Rishindra Reddy, Nallasivam Palanisamy, Andrew C. Chang, Anna Truini, Mauro Truini, Dan R. Robinson, David G. Beer, Arul M. Chinnaiyan. *Transcriptome Meta-Analysis of Lung Cancer Reveals Recurrent Aberrations in NRG1, NF1 and Hippo Pathway Genes*. **(In revision)**

Published/ accepted

1. **Rohit Malik**, Lalit Patel, John R. Prensner, Yang Shi, Matthew Iyer, Shruthi Subramaniyan, Alexander Carley, Yashar S. Niknafs, Anirban Sahu, Sumin Han, Teng Ma, Meilan Liu, Irfan Asangani, Xiaojun Jing, Xuhong Cao, Mohan Dhanasekaran, Dan Robinson, Felix Y. Feng, Arul M. Chinnaiyan. *The lncRNA PCAT29 Inhibits Oncogenic Phenotypes in Prostate Cancer*. **Molecular Cancer Research**. **(Accepted)**
2. Xiaojun Wang, J. Chad Brenner, Irfan A. Asangani, Bushra Ateeq, Yuanyuan Qiao, Marcin Cieslik, Yang Shi, Balabhadrapatruni V. S. K. Chakravarthi, **Rohit Malik**, Xuhong Cao, Xiaojun Jing, Qi Cao, Cynthia X. Wang, Ingrid Apel, Rui Wang, Wei Yan, Hui Jiang, Sooryanarayana Varambally, Shaomeng Wang, and Arul M. Chinnaiyan. *Development of peptidomimetic inhibitors of the ERG transcription factor in prostate cancer*. **Cancer Cell** **(Accepted)**
3. Irfan A. Asangani, Lakshmi Dommeti, Xiaojun Wang, **Rohit Malik**, Yang Rendong, Kari Wilder-Romans, Sudheer Dhanireddy, Mathew K. Iyer, Yi-Mi Wu, Xuhong Cao, Zhaohui S. Qin, Shaomeng Wang, Felix Y. Feng, Arul M. Chinnaiyan. *Therapeutic Targeting of BET Bromodomain Proteins in Castration-Resistant Prostate Cancer*. 2014. **Nature**. 2014 Jun 12;510(7504):278-82
4. John R. Prensner*, Anirban Sahu*, Matthew K. Iyer*, **Rohit Malik***, Benjamin Chandler, Irfan A. Asangani, Anton Poliakov, Ismael A. Vergara, Mohammed Alshalalfa, Robert B. Jenkins, Elai Davicioni, Felix Y. Feng, Arul M. Chinnaiyan. *The lncRNAs PCGEM1 and PRNCR1 are not implicated in castration resistant prostate cancer*. **Oncotargets**. 2014 Mar 30;5(6):1434-8. **(*Co-first Authors)**.
5. John R. Prensner, Wei Chen, Matthew K. Iyer, Qi Cao, Teng Ma, Sumin Han, Anirban Sahu, **Rohit Malik** et.al., *PCAT-1, a long noncoding RNA, regulates BRCA2 and controls homologous recombination in cancer*. **Cancer Research**. 2014 Mar 15;74(6):1651-60
6. Qi Cao, Xiaojun Wang, Meng Zhao, Rendong Yang, **Rohit Malik**, Yuanyuan Qiao, Anton Poliakov, Anastasia K. Yocum, Yong Li, Wei Chen, Xuhong Cao, Xia Jiang, Arun Dahiya, Clair Harris, Felix Y. Feng, Sundeep Kalantry, Zhaohui, S. Qin, Saravana M. Dhanasekaran, and Arul M. Chinnaiyan. *The Central Role of EED in the Orchestration of Polycomb Group Complexes*. 2013. **Nature Communication**. 2014;5:3127
7. O. Alejandro Balbin, John Prensner, Anirban Sahu, Anastasia Yocum, Sunita Shankar, **Rohit Malik**, Damian Fermin, Mohan Dhanasekaran, Benjamin Chandler, Dafydd Thomas, David Beer, Xuhong Cao, Alexey I. Nesvizhskii, and Arul M. Chinnaiyan. *Reconstructing targetable pathways in lung cancer by integrating transcriptome proteome and phosphoproteome*. **Nature Communication**. 2013;4:2617.
8. John R. Prensner, Matthew K. Iyer, Anirban Sahu, Irfan A. Asangani, Qi Cao, Lalit Patel, Ismael A. Vergara, Elai Davicioni, Nicholas Erho, Mercedeh Ghadessi, Robert B. Jenkins, Timothy J. Triche, **Rohit Malik**, et. al., *The lncRNA SchLAP1 coordinates aggressive prostate cancer and antagonizes the SWI/SNF complex*. **Nature Genetics**. 2013; Sep 29. doi: 10.1038/ng.2771

9. **Rohit Malik**, Eunice Soh, JoAnn Trejo and Adriano Marchese. *Novel roles for the E3 ubiquitin ligase AIP4 and STAM-1 in G protein-coupled receptor signaling*. **Journal of Biological Chemistry**, 2013; 287:9013-9027
10. Zannel Blanchard*, **Rohit Malik***, Nicole Mullins, Christine Maric, Hugh Luk, David Horio, Brenda Hernandez, Jeffrey Killeen and Wael M. ElShamy. *Geminin overexpression induces mammary tumors via suppressing cytokinesis*. **Oncotargets**. 2011 Dec;2(12):1011-27 (*Co-first Authors)
11. Lauren Gardner*, **Rohit Malik***, Yoshiko Shimizu, Nicole Mullins and Wael M. ElShamy. *Geminin overexpression prevents the completion of topoisomerase II α chromosome decatenation leading to aneuploidy in human mammary epithelial cells*. **Breast Cancer Research**. 2011 May 19;13(3):R53. (*Co-first Authors)
12. Debra Wyatt*, **Rohit Malik***, Alissa C. Vesecky, and Adriano Marchese. *SUMO modification of arrestin-3 regulates receptor trafficking*. **Journal of Biological Chemistry**. 2011; 286: 3884-3893 (*Co-first Author)
13. **Rohit Malik** and Adriano Marchese. *Arrestin-2 Interacts with the ESCRT machinery to modulate endosomal sorting of CXCR4*. **Molecular Biology of the Cell**. 2010; July 15; 21(14): 2529-41
14. **Rohit Malik** and Sanjay Chhibber. *Protection with bacteriophage K ϕ 1 against fatal Klebsiella pneumoniae-induced burn wound infection in mice*. **Journal of Microbiology Immunology and Infection**. 2009; 42, 134-40
15. Rajiv Giroti, Sunita Verma, Kulwant Singh, **Rohit Malik**, and Indu Talwar. *A grey zone approach for evaluation of 15 short tandem repeat loci in sibship analysis: a pilot study in Indian subjects*. **Journal of Forensic and Legal Medicine**. 2007; 14(5), 261-265

Book Chapter

1. **Rohit Malik** and Adriano Marchese (2012). *Role of β -Arrestins in Endosomal Sorting of G Protein-Coupled Receptors*, **Chemical Biology**, Prof. Deniz Ekinici (Ed.), ISBN: 978-953-51-0049-2, InTech, Available from: <http://www.intechopen.com/books/chemical-biology/role-of-arrestins-in-endosomal-sorting-of-g-protein-coupled-receptors>

Other Media (Non-peer Reviewed)

1. Matthew Iyer, **Rohit Malik**, Anirban Sahu, Javed Siddiqui, Arul Chinnaiyan. *qPCR Validation of Novel Prostate-specific Long Non-coding RNAs*. Application Note; Wafergen Biosciences; 2013. http://www.wafergen.com/wp-content/uploads/2013/01/UM_lncRNA_Tnf.pdf

Conference abstracts

1. **Rohit Malik**, Amjad P. Khan, John R Prensner, Matthew K Iyer, Dmitry Borkin, George Wang, Xia Jiang, Shruthi Subramaniam, Yang Shi, Rachell Stender, YiMi Wu, Xuhong Cao, Jolanta Grembecka, Tomek Cierpicki, Arul M Chinnaiyan. In: Proceedings of the 105th Annual Meeting of the American Association for Cancer Research (AACR)
2. **Rohit Malik**, Matthew K Iyer, Shruthi Subramaniam, Yasuyuki Hosono, Anirban Sahu, Xia Jiang, Yang Shi, Vishal Kothari, Xuhong Cao, Dan Robinson, Saravana M. Dhanasekaran, Felix Y Feng and Arul M Chinnaiyan. Keystone Symposia Conference. "Long Noncoding RNAs: Marching toward Mechanism" 2014

3. **Rohit Malik**, Matthew K. Iyer, John R. Prensner, Lalit Patel, Sumin Han, Wei Chen, Felix Feng, Arul M. Chinnaiyan. In: Proceedings of the 104th Annual Meeting of the American Association for Cancer Research; 2013 Apr 6-10; Washington, DC. Philadelphia (PA): AACR; *Cancer Res* 2013;73 (8 Suppl):Abstract nr 1120. doi:10.1158/1538-7445.AM2013-1120
4. **Rohit Malik**, Matthew K. Iyer, John R. Prensner, Lalit Patel, Sumin Han, Wei Chen, Felix Feng, Arul M. Chinnaiyan. "Sixth Annual Prostate Cancer Program Retreat". March 18-20, 2013
5. **Rohit Malik** and Adriano Marchese (2011). "Experimental Biology" Symposium Session: New Roles for Arrestins in Signaling, Trafficking. April 9-13, Washington DC, USA. *FASEB J* March 17, 2011 25:628.4
6. WM ElShamy, L Gardner, **R Malik**, Y Shimizu, and N Mullins. *Cancer Res.* December 15, 2011; Volume 71, Issue 24, Supplement 3
7. **Rohit Malik** and Adriano Marchese (2010). "ASBMB Special Symposia on Biochemistry and Cell Biology of ESCRTs in Health and Disease". Oct 14-17, Snowbird, UT, USA.
8. **Rohit Malik** and Adriano Marchese (2009). "49th annual meeting of the American society of cell biology". Dec 4-9, San Diego, CA, USA. *Molecular Biology of the Cell*. 20 (Suppl), abstract # 2276/B653.
9. **Rohit Malik** and Sanjay Chhibber (2005). "16th International Phage Biology Meeting". Aug 7-12, Evergreen State College, WI, USA.
10. **Rohit Malik** and Sanjay Chhibber (2005). "7th European Congress of Chemotherapy and Infection", October 19-22, Florence, Italy. *Journal of Chemother*, 17 (Suppl), 3, 7-141.
11. **Rohit Malik** and Sanjay Chhibber (2005). "Advances in Biomedical Sciences" Shivalik Institute of Paramedical Technology (S.I.P.T.), Chandigarh, India.